

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 00:50:11 ; Search time 42.7489 Seconds  
(without alignments)  
3827.903 Million cell updates/sec

Title: US-09-397-967A-17

Perfect score: 25  
Sequence: 1 KDYY 4

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cn2\_1/uspto.spool/US09397967/runat\_03022004\_175631\_28576/app\_query.fasta\_1.1486  
-DB=GenEmbl -OPMT=fastcap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORES=500 -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967 @CGN 1 1 7598 @runat\_03022004\_175631\_28576 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.or.\*  
21: em.ov.\*  
22: em.pat.\*  
23: em.ph.\*  
24: em.pl.\*  
25: em.ro.\*  
26: em.sts.\*  
27: em.un.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.pln.\*  
36: em.htg.rod.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htg.hum.\*  
40: em.htg.mus.\*  
41: em.htg.other.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	25	100.0	30	10	MMF10153
2	25	100.0	36	9	S78945
C	3	25	100.0	47	6 AR289401
4	25	100.0	49	10	MUSIGDQBT
C	5	25	100.0	51	6 AX159117
C	6	25	100.0	51	6 AX159118
C	7	25	100.0	60	9 HOMTCRGAA
C	8	25	100.0	63	6 BD035124
C	9	25	100.0	69	3 MIAARN01
C	10	25	100.0	71	6 AR193244
C	11	25	100.0	77	4 AF339962
C	12	25	100.0	77	4 AF339966
C	13	25	100.0	77	4 AF339972
C	14	25	100.0	81	4 AF339971
C	15	25	100.0	83	4 AF339969
C	16	25	100.0	83	4 AF339970
C	17	25	100.0	84	9 AY256110
18	25	100.0	85	9	AY256163
C	19	25	100.0	88	9 HSPA22C6
C	20	25	100.0	89	4 AF339961
C	21	25	100.0	89	4 AF339963
C	22	25	100.0	89	4 AF339964
C	23	25	100.0	89	4 AF339965
C	24	25	100.0	89	4 AF339967
C	25	25	100.0	89	4 AF339968
C	26	25	100.0	91	8 ATH521235
C	27	25	100.0	94	6 AX522711
C	28	25	100.0	94	8 ATH521186
C	29	25	100.0	96	9 HSPA22C5
C	30	25	100.0	96	11 HSPA22A11
C	31	25	100.0	97	8 ATH523390
C	32	25	100.0	99	6 HSPA22B5
C	33	25	100.0	99	6 AX571616
C	34	25	100.0	106	11 HSC03P01
C	35	25	100.0	107	6 I45687
C	36	25	100.0	107	6 I55779
C	37	25	100.0	108	6 BD093951
C	38	25	100.0	108	6 BD178903
C	39	25	100.0	110	6 BD093953
C	40	25	100.0	110	6 BD178905
C	41	25	100.0	110	9 AF028102
C	42	25	100.0	117	9 AF028094
C	43	25	100.0	122	6 BD043430
C	44	25	100.0	123	6 BD045835
45	25	100.0	124	8	ATH525744

ALIGNMENTS

RESULT 1

```

MMF10153
LOCUS      30 bp      DNA      linear      ROD 07-MAR-1997
DEFINITION M.musculus immunoglobulin heavy chain CDR3 region (focus 1015 DNA
#3).
ACCESSION X67378.1 GI:50904
VERSION    X67378
KEYWORDS   diversity region; immunoglobulin; immunoglobulin heavy chain;
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 30)
AUTHORS    Jacob,J. and Kelsoe,G.
TITLE      In situ studies of the primary immune response to
(4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for
periaarteriolar lymphoid sheath-associated foci and germinal centers
J. Exp. Med. 176 (3), 679-687 (1992)
MEDLINE    92381435
PUBMED     1512536
REFERENCE  2 (bases 1 to 30)
AUTHORS    Jacob,J., Przylepa,J., Miller,C. and Kelsoe,G.
TITLE      In situ studies of the primary immune response to
(4-hydroxy-3-nitrophenyl)acetyl. III. The kinetics of V region
mutation and selection in germinal center B cells
J. Exp. Med. 178 (4), 1293-1307 (1993)
MEDLINE    93389394
PUBMED     8376935
REFERENCE  3 (bases 1 to 30)
AUTHORS    Jacob,J.
TITLE      Direct Submission
JOURNAL    Submitted (14-JUL-1992) J. Jacob, University of Maryland School of
MEDLINE    Baltimore, Dept of Microbiology & Immunology, 655 W Baltimore St,
PUBMED     Baltimore MD 21 201, USA
AUTHORS    See also X67341-7, X67349-91, J00522, J00529-30, J00532-37 &
J00539.
FEATURES   Location/Qualifiers
            source
            1..30
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /cell_line="PALS-associated focus F1015"
            /cell_type="B cells"
            /tissue_type="spleen"
            /dev_stage="day 10"
            <1..>30
            /note="CDR3"
            /codon_start=1
            /protein_id="CAA47790.1"
            /db_xref="GI:1333908"
            /translation="AKDYVGTGVF"
            CDS      8 a      5 c      9 g      8 t
BASE COUNT 8 a      5 c      9 g      8 t
ORIGIN
Alignment Scores:
Pred. No.:      342      Length:      30
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              10      Gaps:      0
US-09-397-967A-17 (1-4) x MMF10153 (1-30)
QY      1 LysAspTyrTyr 4
Db      4 AGGATTACTAC 15
RESULT 2
LOCUS      36 bp      DNA      linear      PRI 07-MAY-1993
DEFINITION T-cell receptor gamma junctional region Vgamma8/J2.3 [human,
S78945
KEYWORDS   diversity region; immunoglobulin; immunoglobulin heavy chain;
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
REFERENCE  1 (bases 1 to 30)
AUTHORS    Jacob,J. and Kelsoe,G.
TITLE      In situ studies of the primary immune response to
(4-hydroxy-3-nitrophenyl)acetyl. II. A common clonal origin for
periaarteriolar lymphoid sheath-associated foci and germinal centers
J. Exp. Med. 176 (3), 679-687 (1992)
MEDLINE    92381435
PUBMED     1512536
REFERENCE  2 (bases 1 to 30)
AUTHORS    Jacob,J., Przylepa,J., Miller,C. and Kelsoe,G.
TITLE      In situ studies of the primary immune response to
(4-hydroxy-3-nitrophenyl)acetyl. III. The kinetics of V region
mutation and selection in germinal center B cells
J. Exp. Med. 178 (4), 1293-1307 (1993)
MEDLINE    93389394
PUBMED     8376935
REFERENCE  3 (bases 1 to 30)
AUTHORS    Jacob,J.
TITLE      Direct Submission
JOURNAL    Submitted (14-JUL-1992) J. Jacob, University of Maryland School of
MEDLINE    Baltimore, Dept of Microbiology & Immunology, 655 W Baltimore St,
PUBMED     Baltimore MD 21 201, USA
AUTHORS    See also X67341-7, X67349-91, J00522, J00529-30, J00532-37 &
J00539.
FEATURES   Location/Qualifiers
            source
            1..30
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /strain="C57BL/6"
            /db_xref="taxon:10090"
            /cell_line="PALS-associated focus F1015"
            /cell_type="B cells"
            /tissue_type="spleen"
            /dev_stage="day 10"
            <1..>30
            /note="CDR3"
            /codon_start=1
            /protein_id="CAA47790.1"
            /db_xref="GI:1333908"
            /translation="AKDYVGTGVF"
            CDS      8 a      5 c      9 g      8 t
BASE COUNT 8 a      5 c      9 g      8 t
ORIGIN
Alignment Scores:
Pred. No.:      342      Length:      30
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              10      Gaps:      0
US-09-397-967A-17 (1-4) x S78945 (1-36)
QY      1 LysAspTyrTyr 4
Db      15 AGGATTACTAT 26
RESULT 3
LOCUS      47 bp      DNA      linear      PAT 12-JUN-2003
DEFINITION Sequence 1136 from patent US 6537751.
ACCESSION  AR289401
VERSION     AR289401.1 GI:31676685
KEYWORDS    Unknown.
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 47)
AUTHORS     Cohen,D., Chumakov,I. and Blumenfeld,M.
TITLE       Biallelic markers for use in constructing a high density
disequilibrium map of the human genome
Patent: US 6537751-A 1136 25-MAR-2003;
JOURNAL     Location/Qualifiers
FEATURES    source
            1..47
            /organism="unknown"
BASE COUNT 15 a      7 c      7 g      17 t      1 others
ORIGIN
Alignment Scores:
Pred. No.:      503      Length:      47
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0

```

```

DB:
US-09-397-967A-17 (1-4) x AR289401 (1-47)
QY
Db
  1 LysAspTyrTyr 4
  47 AAGGATTACTAT 36

RESULT 4
MUSIGDJBUT
LOCUS
DEFINITION Mus musculus rearranged immunoglobulin heavy chain D-J junction
  (clone 04JH4).
ACCESSION L36768
VERSION L36768.1 GI:927162
KEYWORDS diversity region; immunoglobulin heavy chain; joining region;
  rearranged.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 49)
AUTHORS Chowers,Y., Harwood,J., Holtmeier,W., Kagnoff,M.F. and
  Morzycka-Wroblewska,B.
TITLE No title
JOURNAL Unpublished (1994)
COMMENT Original source text: Mus musculus (clone: 04JH4) bone marrow DNA.
FEATURES
  source
    1..49
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /clone="04JH4"
      /tissue_type="bone marrow"
  misc_feature
    1..49
      /note="D-J junction"
BASE COUNT 13 a 6 c 14 g 16 t
ORIGIN
Alignment Scores:
  Pred. No.: 521 Length: 49
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 10 Gaps: 0
US-09-397-967A-17 (1-4) x MUSIGDJBUT (1-49)
QY
Db
  1 LysAspTyrTyr 4
  21 AAGGATTACTAT 32

RESULT 5
AX159117/c
LOCUS
DEFINITION Sequence 2445 from Patent WO0140521.
ACCESSION AX159117
VERSION AX159117.1 GI:14540448
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
  methods of use thereof
JOURNAL Patent: WO 0140521-A 2445 07-JUN-2001;
  Curagen Corporation (US)
FEATURES
  Location/Qualifiers
    source
      1..51
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
  misc_feature
    26
      /note="2 of 2 allelic variants (2445 is other entry)"
      Accession number cg39592883"
BASE COUNT 10 a 12 c 9 g 20 t
ORIGIN
Alignment Scores:
  Pred. No.: 539 Length: 51
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-09-397-967A-17 (1-4) x AX159117 (1-51)
QY
Db
  1 LysAspTyrTyr 4
  44 AAGGATTATTAC 33

RESULT 7
HUMTCRGAA
LOCUS
DEFINITION Human T-cell receptor Ti unproductively rearranged gamma-chain mRNA
  V8-J2-region, partial cds.
ACCESSION M30891
VERSION M30891.1 GI:339401
KEYWORDS J-region; T-cell receptor; V-region; immunoglobulin gamma-chain;
  processed gene.
  
```

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
26
  /note="1 of 2 allelic variants (2446 is other entry)"
  Accession number cg39592883"
BASE COUNT 10 a 11 c 9 g 21 t
ORIGIN
Alignment Scores:
  Pred. No.: 539 Length: 51
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-09-397-967A-17 (1-4) x AX159117 (1-51)
QY
Db
  1 LysAspTyrTyr 4
  44 AAGGATTATTAC 33
  
```

```

RESULT 6
AX159118/c
LOCUS
DEFINITION Sequence 2446 from Patent WO0140521.
ACCESSION AX159118
VERSION AX159118.1 GI:14540449
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Shimkets,R.A. and Leach,M.
TITLE Nucleic acids containing single nucleotide polymorphisms and
  methods of use thereof
JOURNAL Patent: WO 0140521-A 2446 07-JUN-2001;
  Curagen Corporation (US)
FEATURES
  Location/Qualifiers
    source
      1..51
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
  misc_feature
    26
      /note="2 of 2 allelic variants (2445 is other entry)"
      Accession number cg39592883"
BASE COUNT 10 a 12 c 9 g 20 t
ORIGIN
  
```

```

Alignment Scores:
  Pred. No.: 539 Length: 51
  Score: 25.00 Matches: 4
  Percent Similarity: 100.00% Conservative: 0
  Best Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-09-397-967A-17 (1-4) x AX159118 (1-51)
QY
Db
  1 LysAspTyrTyr 4
  44 AAGGATTATTAC 33

RESULT 7
HUMTCRGAA
LOCUS
DEFINITION Human T-cell receptor Ti unproductively rearranged gamma-chain mRNA
  V8-J2-region, partial cds.
ACCESSION M30891
VERSION M30891.1 GI:339401
KEYWORDS J-region; T-cell receptor; V-region; immunoglobulin gamma-chain;
  processed gene.
  
```

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 60)  
 AUTHORS Littman, D.R., Newton, M., Crommie, D., Ang, S.L., Seidman, J.G.,  
 Gettner, S.N., and Weiss, A.  
 TITLE Characterization of an expressed CD3-associated Ti gamma-chain  
 JOURNAL Nature 326 (6108), 85-88 (1987)  
 MEDLINE 87144613  
 PUBMED 3102973  
 COMMENT Original source text: Human leukemic T-cell line PEER, cDNA to  
 mRNA, clone P-gamma-1.  
 FEATURES source  
 Location/Qualifiers  
 1..60  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /map="llq23"  
 1..60  
 /gene="CD3G"  
 <1..>60  
 /gene="CD3G"  
 /note="TCR Ti gamma-chain"  
 /codon\_start=1  
 /protein\_id="AAA61111.1"  
 /db\_xref="GI:339402"  
 /db\_xref="GDB:G00-119-765"  
 /translation="YYCATWKDYKKLFGSGTTL"  
 18..19  
 /gene="CD3G"  
 /organism="Homo sapiens"  
 22..23  
 /gene="CD3G"  
 /organism="Homo sapiens"  
 19 a 11 c 12 g 18 t  
 BASE COUNT  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 619 Length: 60  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-397-967A-17 (1-4) x HUMTCRGAA (1-60)  
 QY 1 LysAspTyrTyr 4  
 Db 19 AAGGATTATTAT 30  
 RESULT 8  
 BD035124/c  
 LOCUS BD035124  
 DEFINITION Sequence tag and encoded human protein.  
 ACCESSION BD035124  
 VERSION BD035124.1 GI:22576866  
 KEYWORDS JP 2001269182-A/11370.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 63)  
 AUTHORS Edwards, J.B.D.M., Duclair, E. and Jordan, J.Y.  
 TITLE Sequence tag and encoded human protein  
 JOURNAL Patent: JP 2001269182-A 11370 02-OCT-2001;  
 GENSEI  
 COMMENT OS Homo sapiens (human)  
 FN JP 2001269182-A/11370  
 PD 02-OCT-2001  
 PF 24-FEB-2000 JP 2000118773

PR 26-FEB-1999 US 60/122487  
 PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCCLAIR, JEAN YVES  
 JORDAN  
 PC C12N15/09, C07K14/435, C07K16/18, C12N1/19, C12N1/21, PC  
 C12N5/10,  
 PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC  
 G06F15/40  
 CC  
 FH Key Location/Qualifiers.  
 1..63  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 18 a 9 c 8 g 28 t  
 BASE COUNT  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 645 Length: 63  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-397-967A-17 (1-4) x BD035124 (1-63)  
 QY 1 LysAspTyrTyr 4  
 Db 19 AAGGATTACTAC 8  
 RESULT 9  
 M1AARN01/c  
 LOCUS M1AARN01  
 DEFINITION Mosquito mitochondrial tRNA-Met.  
 ACCESSION X00599  
 VERSION X00599.1 GI:12627  
 KEYWORDS transfer RNA; transfer RNA-Met.  
 SOURCE mitochondrial Aedes albopictus (Asian tiger mosquito)  
 ORGANISM Aedes albopictus  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Aedes.  
 REFERENCE 1 (bases 1 to 69)  
 AUTHORS Dubin, D.T. and Hsueh, C.C.  
 TITLE Sequence and structure of a methionine transfer RNA from mosquito  
 mitochondria  
 JOURNAL Nucleic Acids Res. 12 (10), 4185-4189 (1984)  
 MEDLINE 84221353  
 COMMENT Data kindly reviewed (15-MAY-1985) by D.T. Dubin.  
 FEATURES source  
 1..69  
 /organism="Aedes albopictus"  
 /organelle="mitochondrion"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:7160"  
 1..69  
 /product="tRNA-Met"  
 /note="tRNA-Met (omitting terminal CAA)"  
 9  
 /note="1-methyladenosine"  
 /mod\_base=mia  
 24  
 /note="pseudouridine"  
 /mod\_base=p  
 25  
 /note="pseudouridine"  
 /mod\_base=p  
 34  
 /note="N-(9-beta-D-ribofuranosyl)purine-6-yl carbamoyl]threonine"  
 /mod\_base=OTHER  
 61  
 /note="pseudouridine"



```
/mod_base=p
62
/note="pseudouridine"
/mod_base=p
25 a 11 c 8 g 25 t
BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 698 Length: 69
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-397-967A-17 (1-4) x MIAARN01 (1-69)

QY 1 LysAspTyrTyr 4
Db 60 AAGGATTACTAC 49

RESULT 10
LOCUS AR193244 71 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 48 from patent US 6346611.
ACCESSION AR193244
VERSION AR193244.1 GI:20239209
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 71)
AUTHORS Pagratis,N., Lochrie,M. and Gold,L.
TITLE High affinity TGF-beta. nucleic acid ligands and inhibitors
JOURNAL Patent: US 6346611-A 48 12-FEB-2002;
FEATURES
source
BASE COUNT 12 a 22 c 16 g 21 t
ORIGIN

Alignment Scores:
Pred. No.: 715 Length: 71
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-397-967A-17 (1-4) x AR193244 (1-71)

QY 1 LysAspTyrTyr 4
Db 26 AAGGATTACTAC 15

RESULT 11
AF339962/c 77 bp DNA linear MAM 12-JUN-2001
LOCUS Puma concolor clone Pco783 locus FCA035 microsatellite sequence.
DEFINITION
ACCESSION AF339962
VERSION AF339962.1 GI:14348718
KEYWORDS
SOURCE Puma concolor (puma)
ORGANISM Puma concolor
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.
REFERENCE 1 (bases 1 to 77)
AUTHORS Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
TITLE Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma concolor)
JOURNAL Mol. Biol. Evol. 18 (6), 1151-1156 (2001)
MEDLINE 21265424
PUBMED 11371606

REFERENCE
AUTHORS Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
TITLE Direct Submission
JOURNAL Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU, 134 Cheatham, Blacksburg, VA 24061-0321, USA
FEATURES
source
/organism="Puma concolor"
/mol_type="genomic DNA"
/db_xref="taxon:9696"
/chromosome="D2"
/map="locus FCA035"
/clone="Pco588"
repeat_region 1..77
/note="microsatellite"
/rpt_type=tandem
19 a 6 c 19 g 33 t
BASE COUNT
ORIGIN

Alignment Scores:
Pred. No.: 766 Length: 77
Score: 25.00 Matches: 4
```

```

Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 4                             Gaps: 0

US-09-397-967A-17 (1-4) x AF339966 (1-77)

QY      1 LysAspTyrTyr 4
DB      47 AAAGATTACTAC 36

RESULT 13
LOCUS   AF339972              77 bp  DNA  linear  MAM 12-JUN-2001
DEFINITION Puma concolor clone Pco201 locus FCA035 microsatellite sequence.
ACCESSION AF339972
VERSION   AF339972.1 GI:14348731
KEYWORDS  Puma concolor (puma)
SOURCE    Puma concolor
ORGANISM  Puma concolor
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.
REFERENCE 1 (bases 1 to 77)
AUTHORS  Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
TITLE     Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma
          concolor)
JOURNAL   Mol. Biol. Evol. 18 (6), 1151-1156 (2001)
MEDLINE   21265424
PUBMED    11371606
REFERENCE 2 (bases 1 to 77)
AUTHORS  Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
TITLE     Direct Submission
JOURNAL   Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU,
          134 Cheatham, Blacksburg, VA 24061-0321, USA
FEATURES  Location/Qualifiers
          source
            1..77
            /organism="Puma concolor"
            /mol_type="genomic DNA"
            /db_xref="taxon:9696"
            /chromosome="D2"
            /map="locus FCA035"
            /clone="Pco201"
          repeat_region
            1..77
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_6_c 19 g 33 t

BASE COUNT 19 a 19 g 33 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
//

Alignment Scores:
Pred. No.: 766      Length: 77
Score: 25.00        Matches: 4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 4                             Gaps: 0

US-09-397-967A-17 (1-4) x AF339972 (1-77)

QY      1 LysAspTyrTyr 4
DB      47 AAAGATTACTAC 36

RESULT 14
LOCUS   AF339971              81 bp  DNA  linear  MAM 12-JUN-2001
DEFINITION Puma concolor clone Pco707 locus FCA035 microsatellite sequence.
ACCESSION AF339971
VERSION   AF339971.1 GI:14348730
KEYWORDS  Puma concolor (puma)
SOURCE    Puma concolor
ORGANISM  Puma concolor
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Puma.
REFERENCE 1 (bases 1 to 81)
AUTHORS  Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
TITLE     Patterns of size homoplasy at 10 microsatellite loci in pumas (Puma
          concolor)
JOURNAL   Mol. Biol. Evol. 18 (6), 1151-1156 (2001)
MEDLINE   21265424
PUBMED    11371606
REFERENCE 2 (bases 1 to 83)
AUTHORS  Culver,M., Menotti-Raymond,M.A. and O'Brien,S.J.
TITLE     Direct Submission
JOURNAL   Submitted (24-JAN-2001) Fisheries and Wildlife Sciences, VPI&SU,
          134 Cheatham, Blacksburg, VA 24061-0321, USA
FEATURES  Location/Qualifiers
          source
            1..83
            /organism="Puma concolor"
            /mol_type="genomic DNA"
            /db_xref="taxon:9696"
            /chromosome="D2"
            /map="locus FCA035"
            /clone="Pco707"
          repeat_region
            1..81
            /note="microsatellite"
            /rpt_type=tandem
            /rpt_6_c 19 a 21 g 35 t

BASE COUNT 19 a 21 g 35 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
23
```

Thu Feb 5 11:06:36 2004

BASE COUNT 19 a /rpt\_type=tandem 22 g 36 t  
ORIGIN

Alignment Scores: 817 Length: 83  
Pred. No.: 25.00 Matches: 4  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 4

US-09-397-967A-17 (1-4) x AF339969 (1-83)

Qy 1 LysAspTyrTyr 4  
Db 47 AAAGATTACTAC 36

Search completed: February 4, 2004, 06:34:16  
Job time : 45.7489 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 4, 2004, 01:30:36 ; Search time 0.62738 Seconds  
(without alignments)  
2814.136 Million cell updates/sec

Title: US-09-397-967A-17

Perfect score: 25

Sequence: 1 KDYY 4

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09397967/runat\_03022004\_175632\_28596/app\_query.fasta\_1.1486  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967 @CGL 1.1.81 @runat\_03022004\_175632\_28596 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTCUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	100.0	29	US-09-304-232-625	Sequence 625, Appl
c	25	100.0	47	US-09-422-978-1136	Sequence 1136, Ap
c	25	100.0	71	US-09-363-939A-48	Sequence 48, Appl
4	25	100.0	107	US-08-441-591-24	Sequence 24, Appl
5	25	100.0	107	US-08-303-362A-24	Sequence 24, Appl
6	25	100.0	107	PCT-US95-05600-41	Sequence 41, Appl
7	25	100.0	192	US-09-107-532A-3404	Sequence 3404, Ap
c	25	100.0	200	US-09-495-050A-146	Sequence 146, Appl
9	25	100.0	270	US-09-328-111-229	Sequence 229, Appl
10	25	100.0	276	US-09-134-001C-744	Sequence 744, Appl
11	25	100.0	282	US-09-313-294A-3240	Sequence 3240, Ap
12	25	100.0	291	US-09-313-294A-340	Sequence 340, Appl

c	13	25	100.0	301	3	US-09-388-917-2	Sequence 2, Appli
	14	25	100.0	301	4	US-09-439-313-250	Sequence 250, App
	15	25	100.0	301	4	US-09-352-616A-250	Sequence 250, App
	16	25	100.0	301	4	US-09-232-149A-250	Sequence 250, App
	17	25	100.0	310	4	US-09-071-035-55	Sequence 55, Appl
	18	25	100.0	311	3	US-09-284-782-29	Sequence 29, Appl
	19	25	100.0	344	2	US-08-975-316-38	Sequence 38, Appl
	20	25	100.0	344	4	US-09-615-192A-38	Sequence 38, Appl
	21	25	100.0	348	2	US-08-883-344-2	Sequence 2, Appli
	22	25	100.0	348	4	US-09-672-609-18	Sequence 18, Appl
	23	25	100.0	348	4	US-09-025-403A-18	Sequence 18, Appl
	24	25	100.0	354	3	US-08-767-128-21	Sequence 21, Appl
	25	25	100.0	357	4	US-09-643-597-196	Sequence 196, App
	26	25	100.0	357	4	US-09-480-884A-196	Sequence 196, App
	27	25	100.0	357	4	US-09-542-615A-196	Sequence 196, App
	28	25	100.0	357	4	US-09-606-421B-196	Sequence 196, App
	29	25	100.0	358	2	US-08-975-316-40	Sequence 40, Appl
	30	25	100.0	358	4	US-09-615-192A-40	Sequence 40, Appl
	31	25	100.0	366	3	US-08-908-643C-54	Sequence 54, Appl
	32	25	100.0	372	4	US-09-672-609-19	Sequence 19, Appl
	33	25	100.0	372	4	US-09-025-403A-19	Sequence 19, Appl
	34	25	100.0	373	1	US-08-017-570-3	Sequence 3, Appli
	35	25	100.0	373	1	US-08-017-570-5	Sequence 5, Appli
	36	25	100.0	373	1	US-08-471-426-3	Sequence 3, Appli
	37	25	100.0	373	1	US-08-471-426-5	Sequence 5, Appli
	38	25	100.0	373	5	PCT-US94-01709-3	Sequence 5, Appli
	39	25	100.0	373	5	PCT-US94-01709-5	Sequence 5, Appli
	40	25	100.0	375	4	US-09-107-532A-1119	Sequence 1119, Ap
c	41	25	100.0	381	4	US-09-091-725-32	Sequence 32, Appl
c	42	25	100.0	384	4	US-09-107-532A-1406	Sequence 1406, Ap
	43	25	100.0	385	3	US-08-732-708C-42	Sequence 42, Appl
	44	25	100.0	390	4	US-09-107-532A-1230	Sequence 1230, Ap
	45	25	100.0	394	4	US-09-242-913B-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-09-304-232-625  
; Sequence 625, Application US/09304232  
; Patent No. 6525185  
; GENERAL INFORMATION:  
; APPLICANT: Fan, Jian Bing  
; APPLICANT: Chakravarti, Aravinda  
; APPLICANT: Halushka, Marc Kenneth  
; APPLICANT: Case Western Reserve University School of Medicine  
; APPLICANT: Affymetrix, Inc.  
; TITLE OF INVENTION: Polymorphisms Associated With  
; FILE REFERENCE: 018547-034210US  
; CURRENT APPLICATION NUMBER: US/09/304,232  
; CURRENT FILING DATE: 1999-05-03  
; EARLIER APPLICATION NUMBER: US 60/084,641  
; EARLIER FILING DATE: 1998-05-07  
; NUMBER OF SEQ ID NOS: 909  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 625  
; LENGTH: 29  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: MREX2 1405  
US-09-304-232-625

Alignment Scores:  
Pred. No.: 60.2 Length: 29  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0  
US-09-397-967A-17 (1-4) x US-09-304-232-625 (1-29)

QY 1 LysAspTyrTyr 4  
| | | | |  
Db 7 AAGACTATAT 18

## RESULT 2

US-09-422-978-1136/c  
; Sequence 1136, Application US/09422978  
; Patent No. 6537751

## GENERAL INFORMATION:

; APPLICANT: Cohen, Daniel  
; APPLICANT: Blumenfeld, Marta  
; APPLICANT: Chumakov, Ilya  
; TITLE OF INVENTION: Biallelic markers for use in constructing a high density...  
; FILE REFERENCE: GENSET 020CP1  
; CURRENT APPLICATION NUMBER: US/09/422,978  
; CURRENT FILING DATE: 1999-10-20  
; EARLIER APPLICATION NUMBER: US 09/298,850  
; EARLIER FILING DATE: 1999-04-21  
; EARLIER APPLICATION NUMBER: US 60/109,732  
; EARLIER FILING DATE: 1998-11-23  
; EARLIER APPLICATION NUMBER: US 60/082,614  
; EARLIER FILING DATE: 1998-04-21  
; NUMBER OF SEQ ID NOS: 11796

SEQ ID NO 1136

LENGTH: 47

TYPE: DNA

ORGANISM: Homo Sapiens

FEATURE:

NAME/KEY: allele

LOCATION: 24

OTHER INFORMATION: 99-20642-382 : polymorphic base A or G

US-09-422-978-1136

## Alignment Scores:

Pred. No.: 96.3 Length: 47  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-422-978-1136 (1-47)

QY 1 LysAspTyrTyr 4  
| | | | |  
Db 47 AAGGATTACTAT 36

## RESULT 3

US-09-363-939A-48/c

; Sequence 48, Application US/09363939A

; Patent No. 6346611

## GENERAL INFORMATION:

; APPLICANT: Pagratis, Nikos  
; APPLICANT: Lochrie, Michael  
; TITLE OF INVENTION: High Affinity TGF $\beta$  Nucleic Acid Ligands and  
; FILE REFERENCE: NEX87  
; CURRENT APPLICATION NUMBER: US/09/363,939A  
; CURRENT FILING DATE: 1999-07-29  
; PRIOR APPLICATION NUMBER: 09/046,247  
; PRIOR FILING DATE: 1998-03-23  
; PRIOR APPLICATION NUMBER: 08/458,424  
; PRIOR FILING DATE: 1995-06-02  
; PRIOR APPLICATION NUMBER: 07/714,131  
; PRIOR FILING DATE: 1991-06-10  
; PRIOR APPLICATION NUMBER: 07/931,473  
; PRIOR FILING DATE: 1992-08-17  
; PRIOR APPLICATION NUMBER: 07/964,624  
; PRIOR FILING DATE: 1992-10-21  
; PRIOR APPLICATION NUMBER: 08/117,991  
; PRIOR FILING DATE: 1993-09-08

; PRIOR APPLICATION NUMBER: 07/536,428  
; PRIOR FILING DATE: 1990-06-11  
; NUMBER OF SEQ ID NOS: 216  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 48

LENGTH: 71

TYPE: RNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: Sequence

NAME/KEY: modified base

LOCATION: (1)..(71)

OTHER INFORMATION: All pyrimidines are 2'P.

US-09-363-939A-48

## Alignment Scores:

Pred. No.: 144 Length: 71  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-363-939A-48 (1-71)

QY 1 LysAspTyrTyr 4

| | | | |

Db 26 AAGGATTACTAC 15

## RESULT 4

US-08-441-591-24

; Sequence 24, Application US/08441591

; Patent No. 5637682

## GENERAL INFORMATION:

; APPLICANT: NIEUWLANDT, D., GOLD, L. AND WECKER, M.

; TITLE OF INVENTION: HIGH-AFFINITY

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS

; TITLE OF INVENTION: TO THE TACHYKININ

; TITLE OF INVENTION: SUBSTANCE P

; NUMBER OF SEQUENCES: 66

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Swanson & Bratschun, L.L.C.

; STREET: 8400 E. Prentice Avenue, Suite 200

; CITY: Englewood

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage

; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/441,591

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/303,362

; FILING DATE: 9-SEPTEMBER-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/714,131

; FILING DATE: 10-JUNE-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/931,473

; FILING DATE: 17-AUGUST-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/117,991

; FILING DATE: 8-SEPTEMBER 1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/536,428

; FILING DATE: 11-JUNE-1990

; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/964,624  
FILING DATE: 21-OCTOBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER: NEX21/C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3333  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-441-591-24

Alignment Scores:  
Pred. No.: 214 Length: 107  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967A-17 (1-4) x US-08-441-591-24 (1-107)

Qy 1 LysAspTyrTyr 4  
Db 70 AAAGAUUAUAC 81

RESULT 5  
US-08-303-362A-24  
Sequence 24, Application US/08303362A  
Patent No. 5648214  
GENERAL INFORMATION:  
APPLICANT: NIEUWLANDT, D., GOLD, L. AND WECKER, M.  
TITLE OF INVENTION: HIGH-AFFINITY  
TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS  
TITLE OF INVENTION: TO THE TACHYKININ  
TITLE OF INVENTION: SUBSTANCE P  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 8400 E. Prentice Avenue, Suite 200  
CITY: Englewood  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/303,362A  
FILING DATE: 9-SEPTEMBER-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/931,473  
FILING DATE: 17-AUGUST-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/117,991  
FILING DATE: 8-SEPTEMBER 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/536,428  
FILING DATE: 11-JUNE-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/964,624  
FILING DATE: 21-OCTOBER-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER: NEX21  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 107  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-303-362A-24

Alignment Scores:  
Pred. No.: 214 Length: 107  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967A-17 (1-4) x US-08-303-362A-24 (1-107)

Qy 1 LysAspTyrTyr 4  
Db 70 AAAGAUUAUAC 81

RESULT 6  
PCT-US95-05600-41  
Sequence 41, Application PC/TUS9505600  
GENERAL INFORMATION:  
APPLICANT: GOLD, LARRY  
APPLICANT: NIEUWLANDT, DAN  
APPLICANT: WECKER, MATTHEW  
APPLICANT: SCHNEIDER, DANIEL J.  
APPLICANT: PEIGON, JULI  
APPLICANT: ALLEN, PATRICK  
APPLICANT: SULLENGER, BRUCE A.  
APPLICANT: DOUDNA, JENNIFER, A.  
TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF  
TITLE OF INVENTION: INSULIN RECEPTOR ANTIBODIES, TACHYKININ SUBSTANCE  
TITLE OF INVENTION: P HIV INTEGRASE AND HIV-1 REVERSE TRANSCRIPTASE  
NUMBER OF SEQUENCES: 239  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 8400 E. Prentice Avenue, Suite 200  
CITY: Englewood  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MG  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/05600  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/238,863  
FILING DATE: 06-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/248,632  
FILING DATE: 24-MAY-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/303,362  
FILING DATE: 09-SEPTEMBER-1994  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/361,795  
FILING DATE: 21-DECEMBER-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/117,991  
FILING DATE: 08-SEPTEMBER-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/931,473  
FILING DATE: 17-AUGUST-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/964,624  
FILING DATE: 21-OCTOBER-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/536,428  
FILING DATE: 11-JUNE-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/714,131  
FILING DATE: 10-JUNE-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/536,428  
FILING DATE: 11-JUNE-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER: NEX17/PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 41:  
LENGTH: 107 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
PCT-US95-05600-41

Alignment Scores: 214 Length: 107  
Pred. No.: 25.00 Matches: 4  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 5

US-09-397-967A-17 (1-4) x PCT-US95-05600-41 (1-107)

QY 1 LysAspTyrTyr 4  
Db 70 AAAGAUAUUUAC 81

RESULT 7  
US-107-532A-3404  
Sequence 3404, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucette-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-JUN-1998

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ariniello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 3404:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 192 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: circular  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (B) LOCATION 1...192  
SEQUENCE DESCRIPTION: SEQ ID NO: 3404:  
US-09-107-532A-3404

Alignment Scores: 379 Length: 192  
Pred. No.: 25.00 Matches: 4  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 100.00% Gaps: 0  
DB: 4

US-09-397-967A-17 (1-4) x US-09-107-532A-3404 (1-192)

QY 1 LysAspTyrTyr 4  
Db 117 AAAGATTATTAT 128

RESULT 8  
US-09-495-050A-146/C  
Sequence 146, Application US/09495050A  
Patent No. 6492505  
GENERAL INFORMATION:  
APPLICANT: Roopa, Reddy  
APPLICANT: Guegler, Karl, J.  
APPLICANT: Au-Young, Janice  
TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P  
FILE REFERENCE: PA-0013 US  
CURRENT APPLICATION NUMBER: US/09/495,050A  
CURRENT FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/118,318  
PRIOR FILING DATE: February 1, 1999  
NUMBER OF SEQ ID NOS: 305  
SOFTWARE: PERL Program  
SEQ ID NO 146  
LENGTH: 200  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. 6492505 1590496CTL  
US-09-495-050A-146

Alignment Scores: 394 Length: 200  
Pred. No.: 25.00 Matches: 4  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Gaps: 0



Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-495-050A-146 (1-200)

Qy 1 LysAspTyrTyr 4

Db 56 AAAGACTATTAT 45

#### RESULT 9

US-09-328-111-229

; Sequence 229, Application US/09328111

; Patent No. 6262333

; GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.

; APPLICANT: Steinmann, Kathleen E.

; APPLICANT: Astle, Jon H.

; APPLICANT: Burgess, Christopher C.

; APPLICANT: Bushnell, Steven E.

; APPLICANT: Carroll III, Eddie

; APPLICANT: Catino, Theodore J.

; APPLICANT: Derti, Adnan

; APPLICANT: Ford, Donna M.

; APPLICANT: Lewis, Marcia E.

; APPLICANT: Monahan, John E.

; APPLICANT: Schlegel, Robert

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; FILE REFERENCE: CCD-257 (US)

; CURRENT APPLICATION NUMBER: US/09/328.111

; CURRENT FILING DATE: 1999-06-08

; EARLIER APPLICATION NUMBER: US 60/088.801

; EARLIER FILING DATE: 1998-06-10

; NUMBER OF SEQ ID NOS: 850

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 229

; LENGTH: 270

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1)..(270)

; OTHER INFORMATION: n = A,T,C or G

US-09-328-111-229

Alignment Scores: 528 Length: 270

Pred. No.: 25.00 Matches: 4

Score: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 3 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-328-111-229 (1-270)

Qy 1 LysAspTyrTyr 4

Db 177 AAAGATTATTAT 188

#### RESULT 10

US-09-134-001C-744

; Sequence 744, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134.001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064.964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055.779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 744

; LENGTH: 276

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-744

Alignment Scores:

Pred. No.: 539 Length: 276

Score: 25.00 Matches: 4

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-134-001C-744 (1-276)

Qy 1 LysAspTyrTyr 4

Db 141 AAAGATTACTAC 152

#### RESULT 11

US-09-313-294A-3240

; Sequence 3240, Application US/09313294A

; Patent No. 6476212

; GENERAL INFORMATION:

; APPLICANT: Lalgudi, Raghunath V.

; APPLICANT: Ito, Laura Y.

; APPLICANT: Sherman, Bradley K.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US

; CURRENT APPLICATION NUMBER: US/09/313.294A

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 7600

; SOFTWARE: PERL Program

; SEQ ID NO 3240

; LENGTH: 282

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. 6476212 7006111167H1

; NAME/KEY: unsure

; LOCATION: 29

; OTHER INFORMATION: a, t, c, g, or other

US-09-313-294A-3240

Alignment Scores:

Pred. No.: 550 Length: 282

Score: 25.00 Matches: 4

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 4 Gaps: 0

US-09-397-967A-17 (1-4) x US-09-313-294A-3240 (1-282)

Qy 1 LysAspTyrTyr 4

Db 11 AAGGATTACTAT 22

#### RESULT 12

US-09-313-294A-340

; Sequence 340, Application US/09313294A

; Patent No. 6476212

; GENERAL INFORMATION:

; APPLICANT: Lalgudi, Raghunath V.

; APPLICANT: Ito, Laura Y.

; APPLICANT: Sherman, Bradley K.

; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR

; FILE REFERENCE: PL-0017 US

; CURRENT APPLICATION NUMBER: US/09/313.294A

```

; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 340
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700549028H1
; NAME/KEY: unsure
; LOCATION: 170, 173, 180
; OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-340

```

```

Alignment Scores:
Pred. No.: 567 Length: 291
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-313-294A-340 (1-291)

```

QY 1 LysAspTyrTyr 4
Db 261 AAAGACTACTAT 272

```

#### RESULT 13

```

US-09-388-917-2/c
; Sequence 2, Application US/09388917
; Patent No. 6258542
; GENERAL INFORMATION:
; APPLICANT: Havashizaki, Yoshihide
; TITLE OF INVENTION: Method for Supporting DNA-Fixation and DNA-Fixed
; FILE REFERENCE: 1794-0120P
; CURRENT APPLICATION NUMBER: US/09/388,917
; EARLIER FILING DATE: 1999-09-02
; EARLIER FILING DATE: 1998-09-04
; EARLIER FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Gallus gallus
US-09-388-917-2

```

```

Alignment Scores:
Pred. No.: 586 Length: 301
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-388-917-2 (1-301)

```

QY 1 LysAspTyrTyr 4
Db 48 AAAGATTATAT 37

```

#### RESULT 14

```

US-09-439-313-250
; Sequence 250, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.

```

```

; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-439-313-250

```

```

Alignment Scores:
Pred. No.: 586 Length: 301
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-439-313-250 (1-301)

```

QY 1 LysAspTyrTyr 4
Db 184 AAAGACTACTAT 195

```

#### RESULT 15

```

US-09-352-616A-250
; Sequence 250, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yuqi
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-352-616A-250

```

```

Alignment Scores:
Pred. No.: 586 Length: 301
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

```

US-09-397-967A-17 (1-4) x US-09-352-616A-250 (1-301)

```

QY 1 LysAspTyrTyr 4
Db 184 AAAGACTACTAT 195

```

Search completed: February 4, 2004, 06:28:05  
Job time : 1.62738 secs

---



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 06:28:12 ; Search time 3.95648 Seconds  
(without alignments)  
3724.149 Million cell updates/sec

Title: US-09-397-967A-17

Perfect score: 25

Sequence: 1 KDYY 4

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 184181367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFPO.spool/US09397967/runat\_03022004\_175633\_28650/app.query.fasta\_1.1486  
-DB=Published Applications NA -QFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09397967@cgn2\_1.1.383 @runat\_03022004\_175633\_28650  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubnpa/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubnpa/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubnpa/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubnpa/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubnpa/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubnpa/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubnpa/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubnpa/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubnpa/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubnpa/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubnpa/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubnpa/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubnpa/US09\_NEW\_PUB.seq2:\*  
14: /cgn2\_6/ptodata/1/pubnpa/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubnpa/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubnpa/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubnpa/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubnpa/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
-----					

	1	25	100.0	25	12	US-10-361-208-442	Sequence 442, App
C	2	25	100.0	25	15	US-10-215-112-8501	Sequence 8501, Ap
C	3	25	100.0	25	15	US-10-098-263B-26149	Sequence 26149, A
	4	25	100.0	25	15	US-10-098-263B-35352	Sequence 35352, A
	5	25	100.0	25	15	US-10-098-263B-43800	Sequence 43800, A
	6	25	100.0	25	15	US-10-098-263B-56890	Sequence 56890, A
	7	25	100.0	25	15	US-10-098-263B-63720	Sequence 63720, A
	8	25	100.0	25	15	US-10-098-263B-75996	Sequence 75996, A
	9	25	100.0	25	15	US-10-098-263B-86117	Sequence 86117, A
10	10	25	100.0	25	15	US-10-098-263B-95846	Sequence 95846, A
	11	25	100.0	25	15	US-10-098-263B-105518	Sequence 105518, A
C	12	25	100.0	25	15	US-10-098-263B-115432	Sequence 115432, A
	13	25	100.0	29	13	US-10-336-638-625	Sequence 625, App
	14	25	100.0	30	12	US-10-361-208-63	Sequence 63, Appl
	15	25	100.0	35	13	US-10-153-244-263	Sequence 263, App
	16	25	100.0	39	12	US-10-361-208-433	Sequence 433, App
	17	25	100.0	39	12	US-10-361-208-437	Sequence 437, App
	18	25	100.0	39	12	US-10-361-208-441	Sequence 441, App
	19	25	100.0	42	12	US-10-361-208-65	Sequence 65, Appl
C	20	25	100.0	42	12	US-10-361-208-66	Sequence 66, Appl
	21	25	100.0	42	12	US-10-361-208-198	Sequence 198, App
C	22	25	100.0	42	12	US-10-361-208-199	Sequence 199, App
C	23	25	100.0	47	12	US-10-349-143-1136	Sequence 1136, Ap
	24	25	100.0	99	10	US-09-960-352-13299	Sequence 13299, A
C	25	25	100.0	109	9	US-09-770-696-802	Sequence 802, App
C	26	25	100.0	120	13	US-10-027-632-51768	Sequence 51768, A
C	27	25	100.0	120	14	US-10-027-632-51768	Sequence 51768, A
	28	25	100.0	131	10	US-09-969-373-450	Sequence 450, App
	29	25	100.0	131	10	US-09-969-373-451	Sequence 451, App
	30	25	100.0	138	13	US-10-091-007-163	Sequence 163, App
C	31	25	100.0	148	10	US-09-969-373-836	Sequence 836, App
	32	25	100.0	150	9	US-09-922-217-750	Sequence 750, App
	33	25	100.0	150	10	US-09-833-263-750	Sequence 750, App
	34	25	100.0	153	14	US-10-025-380-750	Sequence 750, App
	35	25	100.0	153	9	US-09-864-761-26750	Sequence 26750, A
C	36	25	100.0	167	10	US-09-796-692-7365	Sequence 7365, Ap
C	37	25	100.0	167	12	US-10-057-475B-7365	Sequence 7365, Ap
C	38	25	100.0	167	12	US-10-154-884B-7365	Sequence 7365, Ap
C	39	25	100.0	167	15	US-10-040-862-7365	Sequence 7365, Ap
	40	25	100.0	177	12	US-09-864-408A-7645	Sequence 7645, Ap
	41	25	100.0	183	10	US-09-867-701-8586	Sequence 8586, Ap
C	42	25	100.0	184	9	US-09-864-761-29922	Sequence 29922, A
C	43	25	100.0	186	9	US-09-770-696-251	Sequence 251, App
C	44	25	100.0	186	11	US-09-754-853A-136	Sequence 136, App
C	45	25	100.0	189	10	US-09-524-035A-619	Sequence 619, App

#### ALIGNMENTS

RESULT 1  
US-10-361-208-442  
; Sequence 442, Application US/10361208  
; Publication No. US20040009167A1  
; GENERAL INFORMATION:  
; APPLICANT: Rider, Todd H.  
; TITLE OF INVENTION: ANTI-PATHOGEN TREATMENTS  
; FILE REFERENCE: 0050.2041-003  
; CURRENT APPLICATION NUMBER: US/10/361.208  
; PRIOR FILING DATE: 2003-02-07  
; PRIOR APPLICATION NUMBER: US 60/355,359  
; PRIOR FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: US 60/355,022  
; PRIOR FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: US 60/432,386  
; PRIOR FILING DATE: 2002-12-10  
; NUMBER OF SEQ ID NOS: 473  
; SOFTWARE: FAbtSEQ for Windows Version 4.0  
; SEQ ID NO 442  
; LENGTH: 25  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: primer

```

US-10-361-208-442
Alignment Scores:
Pred. No.: 320
Score: 25.00
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 12
Gaps: 0

US-09-397-967A-17 (1-4) x US-10-361-208-442 (1-25)

QY 1 LysAspTyrTyr 4
Db 14 AAGGATTACTAC 25

RESULT 2
US-10-215-112-8501/c
; Sequence 8501, Application US/10215112
; Publication No. US20030082596A1
; GENERAL INFORMATION:
; APPLICANT: Michael Mitmann
; TITLE OF INVENTION: Method of Genetic Analysis of Probes:
; FILE REFERENCE: 3119
; CURRENT APPLICATION NUMBER: US/10/215,112
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 14936
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8501
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Oligonucleotide
US-10-215-112-8501

Alignment Scores:
Pred. No.: 320
Score: 25.00
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 15
Gaps: 0

US-09-397-967A-17 (1-4) x US-10-215-112-8501 (1-25)

QY 1 LysAspTyrTyr 4
Db 13 AAGGATTACTAC 2

RESULT 3
US-10-098-263B-26149/c
; Sequence 26149, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 26149
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-26149

Alignment Scores:
Pred. No.: 320
Score: 25.00
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 15
Gaps: 0

```

```

Score: 25.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 15
Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-26149 (1-25)

QY 1 LysAspTyrTyr 4
Db 24 AAGGATTACTAC 13

RESULT 4
US-10-098-263B-35352
; Sequence 35352, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 35352
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-35352

Alignment Scores:
Pred. No.: 320
Score: 25.00
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 15
Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-35352 (1-25)

QY 1 LysAspTyrTyr 4
Db 3 AAGGATTACTAT 14

RESULT 5
US-10-098-263B-43800
; Sequence 43800, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 43800
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-43800

Alignment Scores:
Pred. No.: 320
Score: 25.00
Length: 25
Matches: 4
Conservative: 0
Best Local Similarity: 100.00%
Query Match: 100.00%
Indels: 0
DB: 15
Gaps: 0

```

```
US-09-397-967A-17 (1-4) x US-10-098-263B-43800 (1-25)
Qy 1 LysAspTyrTyr 4
Db 9 AAGACTACTAC 20
RESULT 6
US-10-098-263B-56890
; Sequence 56890, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 56890
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-56890
Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-09-397-967A-17 (1-4) x US-10-098-263B-56890 (1-25)
Qy 1 LysAspTyrTyr 4
Db 4 AAGACTACTAC 15
RESULT 7
US-10-098-263B-63720
; Sequence 63720, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 63720
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-63720
Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-09-397-967A-17 (1-4) x US-10-098-263B-63720 (1-25)
Qy 1 LysAspTyrTyr 4
Db 1 LysAspTyrTyr 4
Db 7 AAGACTACTAC 18
RESULT 10
US-10-098-263B-95846
; Sequence 95846, Application US/10098263B
```

```
Db 5 AAGACTACTAC 16
RESULT 8
US-10-098-263B-75996
; Sequence 75996, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 75996
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-75996
Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-09-397-967A-17 (1-4) x US-10-098-263B-75996 (1-25)
Qy 1 LysAspTyrTyr 4
Db 6 AAGACTACTAT 17
RESULT 9
US-10-098-263B-86117
; Sequence 86117, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 86117
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-86117
Alignment Scores:
Pred. No.: 320 Length: 25
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-09-397-967A-17 (1-4) x US-10-098-263B-86117 (1-25)
Qy 1 LysAspTyrTyr 4
Db 7 AAGACTACTAT 18
RESULT 10
US-10-098-263B-95846
; Sequence 95846, Application US/10098263B
```

```
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 95846
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-95846

Alignment Scores:
Pred. No.: 320          Length: 25
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-95846 (1-25)

QY 1 LysAspTyrTyr 4
Db 11 AAAGACTACTAT 22

RESULT 11
US-10-098-263B-105518
; Sequence 105518, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 105518
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-105518

Alignment Scores:
Pred. No.: 320          Length: 25
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-105518 (1-25)

QY 1 LysAspTyrTyr 4
Db 3 AGGACTATTAC 14

RESULT 12
US-10-098-263B-115432/c
; Sequence 115432, Application US/10098263B
; Publication No. US20030104410A1
; GENERAL INFORMATION:
; APPLICANT: Mittman, Michael
; TITLE OF INVENTION: Human Microarray
; FILE REFERENCE: 3118.1
```

```
; CURRENT APPLICATION NUMBER: US/10/098,263B
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/276,759
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 131066
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 115432
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-098-263B-115432

Alignment Scores:
Pred. No.: 320          Length: 25
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-098-263B-115432 (1-25)

QY 1 LysAspTyrTyr 4
Db 13 AAAGACTATTAC 2

RESULT 13
US-10-336-638-625
; Sequence 625, Application US/10336638
; Publication No. US20030170699A1
; GENERAL INFORMATION:
; APPLICANT: Fan, Jian Bing
; APPLICANT: Chakravarti, Aravinda
; APPLICANT: Halushka, Marc Kenneth
; APPLICANT: Case Western Reserve University School of Medicine
; APPLICANT: Affymetrix, Inc.
; TITLE OF INVENTION: Polymorphisms Associated With
; FILE REFERENCE: 018547-034210US
; CURRENT APPLICATION NUMBER: US/10/336,638
; CURRENT FILING DATE: 2003-01-02
; PRIOR APPLICATION NUMBER: US/09/304,232
; PRIOR FILING DATE: 1999-05-03
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/084,641
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 909
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 625
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: MRLEX2 1405
US-10-336-638-625

Alignment Scores:
Pred. No.: 374          Length: 29
Score: 25.00           Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13                Gaps: 0

US-09-397-967A-17 (1-4) x US-10-336-638-625 (1-29)

QY 1 LysAspTyrTyr 4
Db 7 AAAGACTAYTAT 18

RESULT 14
US-10-361-208-63
; Sequence 63, Application US/10361208
; Publication No. US20040009167A1
```



; GENERAL INFORMATION:  
; APPLICANT: Rider, Todd H.  
; TITLE OF INVENTION: ANTI-PATHOGEN TREATMENTS  
; FILE REFERENCE: 0050.2041-003  
; CURRENT APPLICATION NUMBER: US/10/361,208  
; CURRENT FILING DATE: 2003-02-07  
; PRIOR APPLICATION NUMBER: US 60/355,359  
; PRIOR FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: US 60/355,022  
; PRIOR FILING DATE: 2002-02-07  
; PRIOR APPLICATION NUMBER: US 60/432,386  
; PRIOR FILING DATE: 2002-12-10  
; NUMBER OF SEQ ID NOS: 473  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 63  
; LENGTH: 30  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: primer  
US-10-361-208-63

Alignment Scores:  
Pred. No.: 388 Length: 30  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967A-17 (1-4) x US-10-361-208-63 (1-30)

Oy 1 LysAspTyrTyr 4  
Db 19 AAAGATTACTAC 30

## RESULT 15

US-10-153-244-263  
; Sequence 263, Application US/10153244  
; Publication No. US20030144191A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL TRP CHANNEL FAMILY MEMBER, TRP-PL  
; FILE REFERENCE: D0144 NP  
; CURRENT APPLICATION NUMBER: US/10/153,244  
; CURRENT FILING DATE: 2002-05-22  
; PRIOR APPLICATION NUMBER: US 60/292,599  
; PRIOR FILING DATE: 2001-05-22  
; PRIOR APPLICATION NUMBER: US 60/362,944  
; PRIOR FILING DATE: 2002-03-08  
; NUMBER OF SEQ ID NOS: 335  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 263  
; LENGTH: 35  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-153-244-263

Alignment Scores:  
Pred. No.: 456 Length: 35  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 13 Gaps: 0

US-09-397-967A-17 (1-4) x US-10-153-244-263 (1-35)

Oy 1 LysAspTyrTyr 4  
Db 20 AAAGATTATTAT 31

Search completed: February 4, 2004, 11:29:00  
Job time : 3.95648 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 02:45:11 ; Search time 23.4633 Seconds  
(without alignments)  
4143.406 Million cell updates/sec

Title: US-09-397-967A-17  
Perfect score: 25  
Sequence: 1 KDXY 4

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ps2n.model\_DEV-xml  
-O=/cgn2\_1/USPFO\_spool/US09397967/runat\_03022004\_175632\_28586/app\_query.fasta\_1.1486  
-DB=EST -QFMT=fastap -SUFFIX=ps2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967@cgn 1\_1\_4514@runat\_03022004\_175632\_28586 -NCPU=6 -ICPU=3  
-NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_plh:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	25	100.0	28	29	CC459528	SALK 1303
C 2	25	100.0	45	12	BJ028703	BJ028703
C 3	25	100.0	48	14	CD457715	FG04d_030
C 4	25	100.0	48	29	BZ765762	BZ765762 SALK_1342
C 5	25	100.0	50	9	AUI07400	AUI07400 AUI07400
C 6	25	100.0	58	9	AA129203	zn36d03.r
C 7	25	100.0	59	9	AA574548	vm29f09.r
C 8	25	100.0	61	9	AA770896	vt13a11.r
C 9	25	100.0	61	9	AA526231	ni75a02.s
C 10	25	100.0	64	9	AA595043	no40e11.s
C 11	25	100.0	65	10	BG656411	ib37b02.x
C 12	25	100.0	67	9	AA388894	mpl3h03.r
C 13	25	100.0	68	28	BH792714	SALK 0649
C 14	25	100.0	69	29	AL755459	Arabidops
C 15	25	100.0	72	29	AL764243	Arabidops
C 16	25	100.0	73	28	AZ440056	2M0230N22
C 17	25	100.0	75	28	AZ774526	2M0004M05
C 18	25	100.0	76	28	BH849410	SALK 0696
C 19	25	100.0	76	28	BH849415	SALK 0696
C 20	25	100.0	76	29	BX536177	Arabidops
C 21	25	100.0	79	12	B1446457	dai88g10
C 22	25	100.0	80	28	BH850632	SALK 0716
C 23	25	100.0	81	9	AA080299	SWB51A031
C 24	25	100.0	82	10	BF648113	NF044D09E
C 25	25	100.0	82	28	BH850750	SALK 0717
C 26	25	100.0	83	13	BH834244	T058G01 P
C 27	25	100.0	85	28	AZ767733	1M0567K11
C 28	25	100.0	87	9	A1537909	tp26f01.x
C 29	25	100.0	94	29	AL952445	Arabidops
C 30	25	100.0	97	14	CD406484	Gm CK3135
C 31	25	100.0	98	28	AZ639208	1M0499T02
C 32	25	100.0	100	9	AA206256	zg54e12.r
C 33	25	100.0	100	14	CA907810	PCS04037
C 34	25	100.0	101	9	AT000602	AT000602
C 35	25	100.0	101	12	BF1322315	xx19d07.Y
C 36	25	100.0	102	10	BF155669	QV4-HT089
C 37	25	100.0	103	13	BQ785901	sa61a07.
C 38	25	100.0	103	28	AZ922716	SLG0t3E07
C 39	25	100.0	104	12	B7430536	B7430536
C 40	25	100.0	104	12	BM874367	laa09g02.
C 41	25	100.0	104	14	CA829282	3529_1.1
C 42	25	100.0	105	10	BE702365	CM4-NN102
C 43	25	100.0	105	29	CC390859	PUHEV29TD
C 44	25	100.0	106	9	A1974023	sd15d03.Y
C 45	25	100.0	106	10	BE344906	945029H02

# ALIGNMENTS

RESULT 1  
CC459528/c  
LOCUS CC459528 28 bp DNA linear GSS 30-MAY-2003  
DEFINITION SALK\_130302.51.20.x Arabidopsis thaliana TDNA insertion lines  
Arabidopsis thaliana genomic clone SALK\_130302.51.20.x, genomic survey sequence.  
ACCESSION CC459528  
VERSION CC459528.1 GI:31225991  
KEYWORDS GSS.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 28)  
 REFERENCE  
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab  
 C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shim,P.,  
 Zimmermann,J. and Ecker,J.R.  
 A Sequence-Indexed Library of Insertion Mutations in the  
 Arabidopsis Genome

UNPUBLISHED  
 CONTACT: Joseph R. Ecker  
 Salk Institute Genomic Analysis Laboratory (SIGNAL)  
 The Salk Institute for Biological Studies  
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 Tel: 858 453 4100 x1752  
 Fax: 858 558 6379  
 Email: ecker@salk.edu

This is single pass sequence recovered from the left border of  
 TDNA.  
 Class: TDNA tagged.  
 Location/Qualifiers  
 1. .28  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone="SALK\_130302.51.20.x"  
 /clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
 /notes="PCR was performed on Arabidopsis thaliana lines  
 each of which contains one or more TDNA insertion  
 elements. The resultant fragment for each line was  
 directly sequenced to determine the genomic sequence at  
 the site of insertion. Details of the protocols used can  
 be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

13 a 2 c 3 g 10 t  
 BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.66e+03 Length: 28  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 29 Gaps: 0

US-09-397-967a-17 (1-4) x CC459528 (1-28)  
 QY 1 LysAspTyrTyr 4  
 Db 21 AAAGATTATTAT 10

RESULT 2  
 BJO28703 45 bp mRNA linear EST 06-DEC-2001  
 LOCUS  
 BJO28703 NIBB Mochii normalized Xenopus neurula library Xenopus  
 laevis cDNA clone XL010a10 5', mRNA sequence.  
 DEFINITION  
 BJO28703  
 VERSION  
 BJO28703.1 GI:17398417  
 KEYWORDS  
 EST.  
 SOURCE  
 Xenopus laevis (African clawed frog)

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 45)  
 REFERENCE  
 Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara  
 Y.

Expressed genes in X. laevis embryo  
 UNPUBLISHED  
 CONTACT: Tadasu Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855

; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 1 (bases 1 to 28)  
 REFERENCE  
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab  
 C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shim,P.,  
 Zimmermann,J. and Ecker,J.R.  
 A Sequence-Indexed Library of Insertion Mutations in the  
 Arabidopsis Genome

UNPUBLISHED  
 CONTACT: Joseph R. Ecker  
 Salk Institute Genomic Analysis Laboratory (SIGNAL)  
 The Salk Institute for Biological Studies  
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 Tel: 858 453 4100 x1752  
 Fax: 858 558 6379  
 Email: ecker@salk.edu

This is single pass sequence recovered from the left border of  
 TDNA.  
 Class: TDNA tagged.  
 Location/Qualifiers  
 1. .28  
 /organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone="SALK\_130302.51.20.x"  
 /clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
 /notes="PCR was performed on Arabidopsis thaliana lines  
 each of which contains one or more TDNA insertion  
 elements. The resultant fragment for each line was  
 directly sequenced to determine the genomic sequence at  
 the site of insertion. Details of the protocols used can  
 be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

13 a 2 c 3 g 10 t  
 BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.66e+03 Length: 28  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 29 Gaps: 0

Email: tshini@genes.nig.ac.jp.  
 Location/Qualifiers  
 1. .45  
 /organism="Xenopus laevis"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:8355"  
 /clone="XL010a10"  
 /tissue type="whole embryo"  
 /dev stage="stage 15"  
 /clone\_lib="NIBB Mochii normalized Xenopus neurula  
 library"

16 a 9 c 6 g 9 t 5 others  
 BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.86e+03 Length: 45  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-09-397-967a-17 (1-4) x BJO28703 (1-45)  
 QY 1 LysAspTyrTyr 4  
 Db 32 AAGGACTATTAT 43

RESULT 3  
 CD457715 48 bp mRNA linear EST 03-JUN-2003  
 LOCUS  
 CD457715 Fg04 AACFC ECORC Fusarium graminearum mycelium grown on wheat heads  
 Gibberella zeae cDNA clone Fg04d\_03o10, mRNA sequence.  
 DEFINITION  
 CD457715  
 VERSION  
 CD457715.1 GI:31372455  
 KEYWORDS  
 EST.  
 SOURCE  
 Gibberella zeae

ORGANISM  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 1 (bases 1 to 48)  
 REFERENCE  
 Ouellet,T., Koul,A., Dan,H., Harris,L.J., Chapados,J., Couroux,P.,  
 De Moors,A., Hattori,J., Lacroix,C., Masotti,M., Robert,L.S., Singh  
 J.A., Sprott,D. and Tinker,N.A.

Fusarium graminearum mycelium grown on wheat heads under high  
 humidity conditions  
 UNPUBLISHED  
 CONTACT: Ouellet, Therese  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-food Canada  
 Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
 CANADA  
 Tel: (613) 759-1658  
 Fax: (613) 759-1701  
 Email: ouellettr@em.agr.ca.

Location/Qualifiers  
 1. .48  
 /organism="Gibberella zeae"  
 /mol\_type="mRNA"  
 /strain="DAOM 180378"  
 /db\_xref="taxon:5518"  
 /clone="Fg04d\_03o10"  
 /tissue type="Mycelium"  
 /dev stage="Asexual"  
 /lab\_host="E. coli"  
 /clone\_lib="Fg04 AACFC ECORC Fusarium graminearum mycelium  
 grown on wheat heads"

/notes="Vector: Bluescript SK+; Site 1: EcoRI; Site 2: XhoI  
 ; Fusarium grown on wheat (cv. Roblin) under high  
 humidity. cDNA made using Stratagene kit."  
 20 a 2 c 9 g 16 t 1 others  
 BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.86e+03 Length: 45  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0

US-09-397-967a-17 (1-4) x BJO28703 (1-45)  
 QY 1 LysAspTyrTyr 4  
 Db 32 AAGGACTATTAT 43

RESULT 3  
 CD457715 48 bp mRNA linear EST 03-JUN-2003  
 LOCUS  
 CD457715 Fg04 AACFC ECORC Fusarium graminearum mycelium grown on wheat heads  
 Gibberella zeae cDNA clone Fg04d\_03o10, mRNA sequence.  
 DEFINITION  
 CD457715  
 VERSION  
 CD457715.1 GI:31372455  
 KEYWORDS  
 EST.  
 SOURCE  
 Gibberella zeae

ORGANISM  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.  
 1 (bases 1 to 48)  
 REFERENCE  
 Ouellet,T., Koul,A., Dan,H., Harris,L.J., Chapados,J., Couroux,P.,  
 De Moors,A., Hattori,J., Lacroix,C., Masotti,M., Robert,L.S., Singh  
 J.A., Sprott,D. and Tinker,N.A.

Fusarium graminearum mycelium grown on wheat heads under high  
 humidity conditions  
 UNPUBLISHED  
 CONTACT: Ouellet, Therese  
 Eastern Cereal and Oilseed Research Centre  
 Agriculture and Agri-food Canada  
 Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6,  
 CANADA  
 Tel: (613) 759-1658  
 Fax: (613) 759-1701  
 Email: ouellettr@em.agr.ca.

```

Alignment Scores:
Pred. No.: 3.08e+03 Length: 48
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-397-967a-17 (1-4) x CD457715 (1-48)

Qy 1 LysAspTyrTyr 4
Db 2 AAAGATTACTAT 13

RESULT 4
LOCUS BZ765762/c 48 bp DNA linear GSS 13-MAR-2003
DEFINITION SALK_134270.26.40.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_134270.26.40.x, genomic
survey sequence.
ACCESSION BZ765762
VERSION BZ765762.1 GI:28938315
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 48)
REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
, C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
, Zimmerman,J. and Ecker,J.R.
AUTHORS A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
JOURNAL Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At2g9130.
Class: TDNA tagged.
FEATURES
source
1..48
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_134270.26.40.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 18 a 12 c 6 g 12 t
ORIGIN

Alignment Scores:
Pred. No.: 3.08e+03 Length: 48
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 29 Gaps: 0

US-09-397-967a-17 (1-4) x BZ765762 (1-48)

```

```

Qy 1 LysAspTyrTyr 4
Db 33 AAAGATTACTAT 22

RESULT 5
LOCUS AU107400 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU107400 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
COL01196, mRNA sequence.
ACCESSION AU107400
VERSION AU107400.1 GI:13556921
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 50)
REFERENCE Suzuki,Y., Taira,H., Tanoda,T., Mizushima-Sugano,J., Sese,J., Hata
,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
,Y., Nakamura,Y., Suyama,A. and Sugano,S.
Diverse transcriptional initiation revealed by fine, large-scale
mapping of mRNA start sites
EMBO Rep. 2 (5), 388-393 (2001)
JOURNAL 21270072
PUBMED 11375929
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
,S. Construction and characterization of a full length-enriched and
a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES
source
1..50
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="COL01196"
/clone_lib="Sugano Homo sapiens cDNA library"
BASE COUNT 14 a 9 c 7 g 20 t
ORIGIN

Alignment Scores:
Pred. No.: 3.22e+03 Length: 50
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-397-967a-17 (1-4) x AU107400 (1-50)

Qy 1 LysAspTyrTyr 4
Db 33 AAAGATTACTAT 44

RESULT 6
LOCUS AA129203 58 bp mRNA linear EST 01-AUG-1997
DEFINITION AA129203 r1 Stratagene endothelial cell 937223 Homo sapiens cDNA
clone IMAGE:549509 5' similar to gb:J03040 SPARC PRECURSOR (HUMAN
), mRNA sequence.
ACCESSION AA129203
VERSION AA129203.1 GI:1689055
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 58)
REFERENCE Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chisoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins

```

M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R. and Marra, M. Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 1475 Std Error: 0.00  
Seq primer: -28M13 rev2 from Amersham  
High quality sequence stop: 1.

FEATURES  
source

```
1..58
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3928026"
/db_xref="taxon:9606"
/clone="IMAGE:549509"
/dev_stage="umbilical vein, 1 passage"
/lab_host="SOLR (kanamycin resistant)"
/clone_lib="Stratagene endothelial cell 937223"
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Umbilical vein endothelial cells, passaged once. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3"
```

BASE COUNT 16 a 13 c 19 g 8 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 3.82e+03 Length: 58  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967a-17 (1-4) x AAL29203 (1-58)

QY 1 LysAspTyrTyr 4  
Db 46 AAGGACTATTAC 57

RESULT 7  
AA574548/c

LOCUS  
DEFINITION  
vm29f09.r1 Knowles Solter mouse blastocyst B1 Mus musculus cDNA  
clone IMAGE:991625 5', mRNA sequence.

ACCESSION  
AA574548  
VERSION  
AA574548.1 GI:2349174

KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
AUTHORS  
1 (bases 1 to 59)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE  
JOURNAL  
COMMENT  
The WashU-HMI Mouse EST Project  
Unpublished  
Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project  
Washington University School  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:563905

Putative full length read  
vector to vector length is 101.

FEATURES  
source

```
1..59
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:991625"
/tissue_type="blastocyst"
/dev_stage="embryo (pre-implantation)"
/lab_host="DH10B"
/clone_lib="Knowles Solter mouse blastocyst B1"
/note="Organ: embryo; Vector: pSPORT; Site 1: NotI; Site 2: Sall; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: Sall(dT): 5'-CGGTGCACGTCGACGTTTTTTTTTTT-3'. cDNAs were cloned into the NotI/Sall sites of a pSPORT vector (Life Technologies). Two different size selections: B1 (larger inserts) and B3."
```

BASE COUNT 21 a 9 c 9 g 20 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.9e+03 Length: 59  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967a-17 (1-4) x AA574548 (1-59)

QY 1 LysAspTyrTyr 4  
Db 33 AAGGACTACTAT 22

RESULT 8  
AA770896

LOCUS  
DEFINITION

AA770896 61 bp mRNA linear EST 29-JAN-1998  
vt13a11.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone  
IMAGE:1162940 5', similar to SW:KGUA\_MOUSE Q64520 GUANYLATE KINASE  
; mRNA sequence.

ACCESSION  
AA770896  
VERSION  
AA770896.1 GI:2822707

KEYWORDS  
EST.  
SOURCE  
Mus musculus (house mouse)

ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
AUTHORS  
1 (bases 1 to 61)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE  
JOURNAL  
COMMENT  
The WashU-HMI Mouse EST Project  
Unpublished  
Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project  
Washington University School  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: mouseest@watson.wustl.edu  
This clone is available royalty-free through LLNL ; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
MGI:628852

Trace considered overall poor quality

Possible reversed clone; similarity on wrong strand

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 1.

#### FEATURES

source

1. .61  
Location/Qualifiers

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C3H"  
/db\_xref="taxon:10090"  
/clone="IMAGE:1162940"  
/cell\_lines="C2C12"  
/lab\_host="DH10B"  
/clone\_lib="Barstead mouse myotubes MFLRB5"  
/note="Vector: pT7T3-Pac (Pharmacia) with a modified  
polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5',  
TGTTACGAATCGAAGCGGAGCGCCCTTTTCTTTTCTTTTCTTTTCTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
[AATTCGATCCTTG], digested with Not I and cloned into the  
Not I and Eco RI sites of the modified pT7T3 vector.  
Library constructed by Bob Barstead. The C2C12 cell line  
(available from ATCC, catalog # CRL-1772) differentiates  
rapidly, forming contractile myotubes and producing  
characteristic muscle proteins. "

BASE COUNT 19 a 9 c 18 g 15 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.05e+03 Length: 61  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967a-17 (1-4) x AA770896 (1-61)

Qy 1 LysAspTyrTyr 4  
Db 14 AAGATTACTAC 25

#### RESULT 9

AA526231

LOCUS

DEFINITION

ni75a02.s1 NCI CGAP Pr12 Homo sapiens cDNA clone IMAGE:982634

similar to SW:NUJM\_HUMAN P03897 NADH-UBIQUINONE OXIDOREDUCTASE

CHAIN 3 ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: W. Douglas Figg, Ph.D., Paul H. Duray, M.D.,  
Rodrigo F. Chuqui, M.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 827 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 1.

#### FEATURES

source

1. .61  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:982634"  
/sex="male"  
/tissue\_type="metastatic prostate bone lesion"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr12"  
/note="Vector: pAMP10; mRNA made from metastatic prostate  
lesion of the bone, cDNA made by oligo-dT priming.  
Non-directionally cloned. Size-selected on agarose gel,  
average insert size 600 bp. Library made by D. Krizman,  
NIH."

BASE COUNT 23 a 19 c 8 g 11 t  
ORIGIN

Alignment Scores:

Pred. No.: 4.05e+03 Length: 61  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967a-17 (1-4) x AA526231 (1-61)

Qy 1 LysAspTyrTyr 4

Db 4 AAGGATTATTAC 15

#### RESULT 10

AA595043

LOCUS

DEFINITION

AA595043 no40e11.s1 NCI CGAP Pr23 Homo sapiens cDNA clone IMAGE:1103180 3',  
similar to SW:NU4M\_HUMAN P03905 NADH-UBIQUINONE OXIDOREDUCTASE

CHAIN 4 ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Stratagene, Inc.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality

Insert Length: 1294 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 1.

#### FEATURES

source

1. .64  
Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1103180"
/sex="male, pooled"
/tissue_type="prostate tumor"
/lab_host="NCI CGAP Pr23"
/note="Organ: prostate; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
5' adaptor sequence: 5'
GAATTCGGACGAG 3' 3' adaptor sequence: 5'
CTCGAGTTTTTTTTTTTTT 3' Average insert size: 1.2 kb."
BASE COUNT      22 a      7 c      19 g      16 t
ORIGIN

Alignment Scores:
Pred. No.:      4.28e+03      Length:      64
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              9      Gaps:      0

US-09-397-967A-17 (1-4) x AA595043 (1-64)

QY      1 LysAspTyrTyr 4
Db      41 AAGGATTATTAT 52

RESULT 11
BG656411/c
LOCUS      BG656411      65 bp      mRNA      linear      EST 05-JUL-2001
DEFINITION      ib37b02.x1 HR85 islet Homo sapiens cDNA 3', mRNA sequence.
ACCESSION      BG656411
VERSION      BG656411.1      GI:13793820
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 65)
Melton, D., Brown, J., Kent, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, J., Clifton, S.,
Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, B., Ronko, I., Bennett, J., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T.,
Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Other ESPs: ib37b02.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Harvard Medical School
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohpc.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .65
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:

```

```

NotI; Site 2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size ~1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:
314-362-1916, Fax: 314-747-2692."
BASE COUNT      25 a      5 c      9 g      26 t
ORIGIN

Alignment Scores:
Pred. No.:      4.36e+03      Length:      65
Score:          25.00      Matches:      4
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              10      Gaps:      0

US-09-397-967A-17 (1-4) x BG656411 (1-65)

QY      1 LysAspTyrTyr 4
Db      14 AAGGATTATTAT 3

RESULT 12
AA388894
LOCUS      AA388894      67 bp      mRNA      linear      EST 23-APR-1997
DEFINITION      mp13h03.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
clone IMAGE:569141 5', similar to TR:G1235727 G1235727 SUPPRESSOR
FOR YEAST MUTANT. ;, mRNA sequence.
ACCESSION      AA388894
VERSION      AA388894.1      GI:2041848
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 67)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:343789
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. .67
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:569141"
/tissue_type="embryo"
/dev_stage="8.5dpc embryos"
/lab_host="DH10B"
/clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
SalI; Site_2: NotI; Cloned unidirectionally. Primer:

```



BASE COUNT 20 a 21 c 11 g 15 t

ORIGIN

Alignment Scores:

Pred. No.: 4.51e+03 Length: 67  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-397-967A-17 (1-4) x AA388894 (1-67)

Qy 1 LysAspTyrTyr 4

Db 26 AAAGACTACTAC 37

RESULT 13

LOCUS BH792714

DEFINITION BH792714 68 bp DNA linear GSS 02-APR-2002  
Arabidopsis thaliana genomic clone SALK\_064938.46.55.x, genomic survey sequence.

ACCESSION BH792714

VERSION BH792714.1

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 68)  
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.

TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL Unpublished

COMMENT Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of TDNA.

FEATURES

source

Location/Qualifiers

1..68  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="SALK\_064938.46.55.x"  
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at [http://signal.salk.edu/tdna\\_protocols.html](http://signal.salk.edu/tdna_protocols.html)"

BASE COUNT 19 a 11 c 9 g 29 t

ORIGIN

Alignment Scores:

Pred. No.: 4.59e+03 Length: 68  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 28 Gaps: 0

US-09-397-967A-17 (1-4) x BH792714 (1-68)

Qy 1 LysAspTyrTyr 4

Db 13 AAGGATTATTAT 24

RESULT 14

LOCUS AL755459/c

DEFINITION AL755459 69 bp DNA linear GSS 17-JUN-2002  
Arabidopsis thaliana T-DNA flanking sequence GK-098A06-012117, genomic survey sequence.

ACCESSION AL755459

VERSION AL755459.1

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 69)  
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoveer,P., Dekker,K., Saedler,H. and Weisshaar,B.  
TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines

JOURNAL Unpublished

REFERENCE 2

A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics

JOURNAL Unpublished

REFERENCE 3

(bases 1 to 69)

Li,Y., Rosso,M., Strizhov,N. and Weisshaar,B.  
Direct Submission  
Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

JOURNAL This sequence is recovered from the right border of the T-DNA. It indicates an insertion within the locus defined by clone F12F1. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at: <http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

FEATURES

source

Location/Qualifiers

1..69  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="GK-098A06-012117"  
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector PAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 24 a 13 c 5 g 27 t

ORIGIN

Alignment Scores:

Pred. No.: 4.67e+03 Length: 69  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 29 Gaps: 0

US-09-397-967A-17 (1-4) x AL755459 (1-69)

QY 1 LysAspTyrTyr 4  
 |||||  
 Db 16 AAAGATTATTAT 5

Db 54 AAAGATTACTAC 65  
 |||||

Search completed: February 4, 2004, 08:14:56  
 Job time : 29.4633 secs

RESULT 15  
 AL764243 72 bp DNA linear GSS 18-JUN-2002  
 LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-123E01-012812,  
 DEFINITION genomic survey sequence.

ACCESSION AL764243  
 VERSION AL764243.1 GI:21516777  
 KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)  
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Siedler, H.  
 and Weisshaar, B.

TITLE A pipeline for automated high-throughput generation of FSTs  
 (flanking sequence tags) from Arabidopsis thaliana T-DNA

JOURNAL transformed lines  
 REFERENCE 2 Unpublished

AUTHORS Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.  
 TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)  
 for flanking sequence tag based reverse genetics

JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 72)

AUTHORS Rosso, M., Li, Y., Strizhov, N. and Weisshaar, B.

TITLE Direct Submission  
 JOURNAL Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer  
 Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany  
 COMMENT This sequence is recovered from the left border of the T-DNA. It  
 indicates an insertion close to or within gene At3g05630. The  
 sequences are generated at the MPI for Plant Breeding Research in  
 the context of the GABI-Kat project. GABI-Kat is part of the German  
 Plant Genomics program designated 'GABI'. Information on line  
 availability can be found at:  
 http://www.mpiz-koeln.mpg.de/GABI-Kat/.

FEATURES  
 1. .72 Location/Qualifiers

/organism="Arabidopsis thaliana"  
 /mol\_type="genomic DNA"  
 /strain="Columbia 0"  
 /db\_xref="taxon:3702"  
 /clone="GK-123E01-012812"  
 /notes="PCR was performed on DNA from Arabidopsis thaliana  
 plants (T1) which were transformed with the T-DNA from  
 vector pAC161. The lines contain one or more T-DNA  
 insertions. The DNA fragment(s) resulting from the PCR  
 were directly sequenced to determine the genomic sequence  
 flanking the insertion. Sequences displaying significant  
 similarity to the A. thaliana nuclear genome sequence were  
 processed for submission. T-DNA derived sequences were  
 removed"

BASE COUNT 24 a 16 c 15 g 17 t  
 ORIGIN

Alignment Scores:

Pred. No.:	4.9e+03	Length:	72
Score:	25.00	Matches:	4
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	29	Gaps:	0

US-09-397-967A-17 (1-4) x AL764243 (1-72)

QY 1 LysAspTyrTyr 4

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 00:42:56 ; Search time 2.90481 Seconds  
(without alignments)  
3717.204 Million cell updates/sec

Title: US-09-397-967A-17  
Perfect score: 25  
Sequence: 1 KDYY 4

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09397967/runat\_03022004.175631.28566/app\_query.fasta\_1.1486  
-DB=N Geneseq 19Jun03 -OFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=100 -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptn -NORM=ext -HRAPISE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967 @CGN 1.1.501 @runat\_03022004.175631.28568 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=10 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	25	100.0	12	23	ABI57762
C 2	25	100.0	12	23	ABI78930
C 3	25	100.0	13	23	ABC49776
C 4	25	100.0	13	23	ABC49777
C 5	25	100.0	22	16	AAQ85884
C 6	25	100.0	24	22	AAI69602
C 7	25	100.0	25	21	AAA68498
C 8	25	100.0	29	21	AAQ44425
C 9	25	100.0	30	24	AAQ69473
C 10	25	100.0	47	21	AAZ66789
C 11	25	100.0	51	22	AAI75504
C 12	25	100.0	51	22	AAI75505
C 13	25	100.0	63	21	AAI1379
C 14	25	100.0	64	13	AAQ33504
C 15	25	100.0	71	22	AAQ33504
C 16	25	100.0	76	21	AAQ60873
C 17	25	100.0	94	25	ABZ78674
C 18	25	100.0	94	25	ABZ09221
C 19	25	100.0	99	25	ABX08127
C 20	25	100.0	99	25	ABX48134
C 21	25	100.0	107	16	AAQ6096
C 22	25	100.0	108	20	AAZ33023
C 23	25	100.0	108	22	AAF87848
C 24	25	100.0	110	20	AAZ33025
C 25	25	100.0	110	22	AAF87850
C 26	25	100.0	119	22	AAK70919
C 27	25	100.0	122	21	AAQ19885
C 28	25	100.0	123	21	AAQ22090
C 29	25	100.0	138	22	AAQ07099
C 30	25	100.0	150	22	AAI29196
C 31	25	100.0	150	25	ABZ33382
C 32	25	100.0	153	22	ABA66324
C 33	25	100.0	153	22	AAI4742
C 34	25	100.0	153	22	AAQ40544
C 35	25	100.0	153	25	ACR00545
C 36	25	100.0	159	22	AAK74089
C 37	25	100.0	165	16	AAQ87512
C 38	25	100.0	177	21	AAQ21103
C 39	25	100.0	177	24	ABN78876
C 40	25	100.0	183	24	ABL85608
C 41	25	100.0	184	22	AAK23938
C 42	25	100.0	184	22	AAK50009
C 43	25	100.0	184	24	ABQ23501
C 44	25	100.0	185	21	AAQ22811
C 45	25	100.0	186	22	AAI61505

ALIGNMENTS

RESULT 1  
ABI57762/c  
ID ABI57762 standard; DNA; 12 BP.  
XX  
XX ABI57762;  
XX  
XX 22-FEB-2002 (first entry)  
XX  
XX Oligonucleotide primer SEQ ID NO 357735 for detecting SNP TSC0050754.  
DE  
XX SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.  
XX  
XX Homo sapiens.  
XX  
XX WO200177384-A2.  
PN

```

XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-IB00713.
PF
XX
XX 07-APR-2000; 2000DE-1019173.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2001-657177/75.
DR
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
XX Claim 1; SEQ ID 357735; 29pp + Sequence Listing; German.
PS
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 12 BP; 5 A; 2 C; 0 G; 5 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 127 Length: 12
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 23 Gaps: 0
US-09-397-967A-17 (1-4) x ABI57762 (1-12)
QV 1 LysAspTyrTyr 4
Db 12 AAGGATTATTAT 1
RESULT 2
ABI78930/c
ID ABI78930 standard; DNA; 12 BP.
XX
XX AC ABI78930;
XX
XX 22-FEB-2002 (first entry)
DT
XX
XX Homo sapiens.
OS
XX
XX WO200177384-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
OS
XX
XX WO200177384-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-IB00713.
PF
XX
XX 07-APR-2000; 2000DE-1019173.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA

```

```

XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2001-657177/75.
DR
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
XX Claim 1; SEQ ID 378903; 29pp + Sequence Listing; German.
PS
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC ABI00010-ABI82073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 12 BP; 5 A; 1 C; 0 G; 6 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 127 Length: 12
Score: 25.00 Matches: 4
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 23 Gaps: 0
US-09-397-967A-17 (1-4) x ABI78930 (1-12)
QV 1 LysAspTyrTyr 4
Db 12 AAGGATTATTAT 1
RESULT 3
ABC49776
ID ABC49776 standard; DNA; 13 BP.
XX
XX AC ABC49776;
XX
XX 21-FEB-2002 (first entry)
DT
XX
XX Oligonucleotide SEQ ID NO 49793 for detecting SNP TSC0014050.
DE
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
OS
XX
XX WO200177384-A2.
PN
XX
XX 18-OCT-2001.
PD
XX
XX 06-APR-2001; 2001WO-IB00713.
PF
XX
XX 07-APR-2000; 2000DE-1019173.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2001-657177/75.
DR
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -

```

XX PS Claim 1; SEQ ID 49793; 29pp + Sequence Listing; German.

XX CC This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The CC CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a CC CC range of diseases including immune system, gastrointestinal, respiratory, CC CC central nervous system, cardiovascular and metabolic disorders. The CC CC oligomers are also used for detecting cell type differentiation. CC CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and CC CC ABI00010-ABI82073 represent the oligomers described in the invention. CC CC NOTE: The sequence data for this patent did not form part of the printed CC CC specification, but was obtained in electronic format from WIPO at CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 13 BP; 7 A; 0 C; 1 G; 5 T; 0 other;

Alignment Scores: Pred. No.: 137 Length: 13  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 23 Gaps: 0

US-09-397-967A-17 (1-4) x ABC49776 (1-13)

QY 1 LysAspTyrTyr 4  
Db 1 AAAGATTATTAT 12

RESULT 4

ABC49777/c ID ABC49777 standard; DNA; 13 BP.

XX AC ABC49777;

XX DT 21-FEB-2002 (first entry)

XX DE Oligonucleotide SEQ ID NO 49794 for detecting SNP TSC0014050.  
XX KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;  
XX KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;  
XX KW central nervous system; gastrointestinal; respiratory; immune; metabolic.

XX OS Homo sapiens.

XX PN WO200177384-A2.

XX XX 18-OCT-2001.

XX PF 06-APR-2001; 2001WO-IB00713.

XX PR 07-APR-2000; 2000DE-1019173.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX XX WPI; 2001-657177/75.

XX DR

XX PT Set of oligonucleotides, useful for diagnosis and cell typing, is  
XX PT designed to detect single nucleotide polymorphisms and cytosine  
XX PT methylation status -  
XX XX  
XX PS Claim 1; SEQ ID 49794; 29pp + Sequence Listing; German.

XX CC This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The CC CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a CC CC range of diseases including immune system, gastrointestinal, respiratory,

CC CC central nervous system, cardiovascular and metabolic disorders. The CC CC oligomers are also used for detecting cell type differentiation. CC CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and CC CC ABI00010-ABI82073 represent the oligomers described in the invention. CC CC NOTE: The sequence data for this patent did not form part of the printed CC CC specification, but was obtained in electronic format from WIPO at CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 13 BP; 5 A; 1 C; 0 G; 7 T; 0 other;

Alignment Scores: Pred. No.: 137 Length: 13  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 23 Gaps: 0

US-09-397-967A-17 (1-4) x ABC49777 (1-13)

QY 1 LysAspTyrTyr 4  
Db 13 AAAGATTATTAT 2

RESULT 5

AAQ85884

ID AAQ85884 standard; CDNA; 22 BP.

XX AC AAQ85884;

XX XX 25-MAR-2003 (updated)

XX DT 24-OCT-1995 (first entry)

XX DE Primer for amplifying gag gene of feline immunodeficiency virus.  
XX KW Feline immunodeficiency virus; FIV; CD4; CD8; antigen; vaccine;  
XX KW diagnosis; primer; probe; acquired immune deficiency syndrome; AIDS;  
XX KW model system; screening; long terminal repeat; LTR; ss.  
XX OS Synthetic.

XX PN WO9505460-A2.

XX XX 23-FEB-1995.

XX PF 25-JUL-1994; 94WO-US08364.

XX PR 12-AUG-1993; 93US-0105710.

XX PR 03-SEP-1991; 91US-0752424.

XX PA (UYNC-) UNIV NORTH CAROLINA STATE.

XX XX Tompkins MB, Tompkins WAF;

XX XX WPI; 1995-098760/13.

XX PT New isolate of feline immunodeficiency virus - and infected host  
XX PT cells, DNA, vectors and infected cats, useful in vaccines and as  
XX PT models for human AIDS  
XX XX  
XX PS Example 12; Page 33; 50pp; English.

XX CC The NCSU strain (ATCC VR2333) of the feline immunodeficiency virus (FIV) is highly infectious in vivo and causes a rapid inversion of the CC CC CD4+CD8+ ratio in infected animals. Antigenic fragments of the CC CC virus can be used in vaccines. Fragments of the DNA can also be used CC CC diagnostically as probes and primers and for expressing viral CC CC antigens. Cats infected with the virus are useful as model systems CC CC for studying AIDS and for screening therapeutic drugs. Two primers CC CC (AAQ85883, AAQ85884) were used to amplify the gag gene of the NCSU1 CC CC strain of FIV (AAQ85885) for its subcloning and subsequent CC CC sequencing. NOTE: This sequence is the primer cited in the CC CC example although a second SEQ ID 5 appears in the sequence listing of

CC the specification. The second sequence is TCTGGCCTCA TTCTAATGA TG,  
 CC this sequences complement.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX

SQ Sequence 22 BP; 10 A; 4 C; 5 G; 3 T; 0 other;

Alignment Scores:  
 Pred. No.: 229 Length: 22  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 16 Gaps: 0

US-09-397-967A-17 (1-4) x AAQ85884 (1-22)

QY 1 LysAspTyrTyr 4  
 Db 11 AAAGATTACTAC 22

RESULT 6  
 AAI69602/C  
 ID AAI69602 standard; DNA; 24 BP.  
 XX  
 AC AAI69602;  
 XX  
 DT 19-DEC-2001 (first entry)  
 DE Human calthrin light chain 17 coding sequence PCR primer #2.  
 XX  
 XX Human; calthrin light chain 17; cancer; haemopathy; HIV infection;  
 KW immunological disease; inflammation; gene therapy; PCR primer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175045-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 26-MAR-2001; 2001WO-CN00456.  
 XX  
 PR 28-MAR-2000; 2000CN-0115248.  
 XX

PA (SHAN-) SHANGHAI BIOWINDOW GENE DEV INC.  
 PI Mao Y, Xie Y;  
 XX  
 XX WPI; 2001-639358/73.  
 XX

PT New human clathrin light chain 17 for diagnosing and treating malignant  
 PT tumor, hemopathy, human immunodeficiency virus infection, immunological  
 PT diseases and inflammation -  
 XX  
 XX Example 2; Page 16; 32pp; Chinese.  
 PS

CC The present invention provides the protein and coding sequences of human  
 CC calthrin light chain 17. The sequences can be used in the treatment of  
 CC cancer, haemopathy, HIV infection, immunological diseases and  
 CC inflammation. The present sequence is a PCR primer for the coding  
 CC sequence of the invention.  
 XX

SQ Sequence 24 BP; 7 A; 2 C; 2 G; 13 T; 0 other;

Alignment Scores:  
 Pred. No.: 249 Length: 24  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-397-967A-17 (1-4) x AAI69602 (1-24)

QY 1 LysAspTyrTyr 4  
 Db 14 AAAGATTATTAC 3

RESULT 7  
 AAA68498  
 ID AAA68498 standard; DNA; 25 BP.  
 XX  
 AC AAA68498;  
 XX  
 DT 27-OCT-2000 (first entry)  
 DE Bacteriophage 3A ORF RBS sequence 3AORF244.  
 XX  
 XX Bacteriophage; antimicrobial; genome; identification; antibacterial;  
 KW bacterial growth inhibition; PCR primer; RBS; ribosome binding site;  
 KW bacterial infection; ss.  
 XX  
 OS Bacteriophage 3A.  
 XX  
 PN WO200032825-A2.  
 XX  
 PD 08-JUN-2000.  
 XX  
 PF 03-DEC-1999; 99WO-IB02040.  
 XX  
 PR 03-DEC-1998; 98US-0110992.  
 PR 03-JUN-1999; 99US-0328144.  
 PR 28-SEP-1999; 99US-0407804.  
 PR 30-SEP-1999; 99US-0157218.  
 PR 01-DEC-1999; 99US-0168777.  
 PR 02-DEC-1999; 99US-0454252.  
 XX  
 PA (PHAG-) PHAGETECH INC.  
 XX  
 PI Pelletier J, Gros P, Dubow M;  
 XX  
 XX WPI; 2000-412361/35.  
 XX

PT Identifying a bacteriophage coding region for treating bacterial  
 PT infections comprises identifying a nucleic acid encoding a product that  
 PT inhibits bacteria when a bacteriophage infects a bacterium -  
 XX  
 XX Disclosure; Page 187; 456pp; English.  
 PS

CC The present invention describes a method for identifying a bacteriophage  
 CC coding region encoding a product active on an essential bacterial  
 CC target. The method comprises identifying a nucleic acid sequence encoding  
 CC a gene product that provides a bacteria-inhibiting function when an  
 CC uncharacterised bacteriophage infects a pathogenic bacterium. The  
 CC compound active on a target of a bacteriophage inhibitor protein in a  
 CC bacteria is used to treat or prevent a bacterial infection in an animal.  
 CC AAA68243 to AAA69442 and AAB16523 to AAB16954 represent bacteriophage  
 CC nucleotide and protein sequences which are used in the exemplification of  
 CC the present invention.  
 XX

SQ Sequence 25 BP; 10 A; 2 C; 2 G; 11 T; 0 other;

Alignment Scores:  
 Pred. No.: 259 Length: 25  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0

US-09-397-967A-17 (1-4) x AAA68498 (1-25)

QY 1 LysAspTyrTyr 4  
 Db 11 AAAGATTATTAT 22

RESULT 8

AAA04425  
ID AAA04425 standard; DNA; 29 BP.  
XX AC AAA04425;  
XX DT 22-MAY-2000 (first entry)  
XX XX Polymorphic fragment of hypertension associated gene MRL.  
XX DE  
XX KW Polymorphism; hypertension; agammaglobulinemia; diabetes insipidus;  
XX Lesch-Nyhan syndrome; muscular dystrophy; Wiskott-Aldrich syndrome;  
XX Fabry's disease; familial hypercholesterolemia; hereditary spherocytosis;  
XX polycystic kidney disease; von Willebrand's disease; forensic; human;  
XX tubercous sclerosis; hereditary hemorrhagica telangiectasia;  
XX familial colonic polyposis; hereditary hemorrhagica telangiectasia;  
XX Ehlers-Danlos syndrome; ss.  
XX OS Homo sapiens.  
XX XX  
XX PN EP955382-A2.  
XX XX  
XX PD 10-NOV-1999.  
XX XX  
XX PF 07-MAY-1999; 99EP-0250150.  
XX XX  
XX PR 07-MAY-1998; 98US-0084641.  
XX PR 03-MAY-1999; 99US-0304232.  
XX XX  
XX PA (AFFY-) AFFYMETRIX INC.  
XX PA (UYCA-) UNIV CASE WESTERN RESERVE.  
XX PI Fan JB, Chakravarti A, Haluska MK;  
XX XX  
XX DR WPI; 2000-107928/10.  
XX XX  
XX PT Novel nucleic acids containing polymorphisms used in the diagnosis of  
XX PT hypertension -  
XX XX  
XX PS Claim 1; Page 36; 53pp; English.  
XX CC The invention provides polymorphic fragments of genes associated with  
XX CC hypertension. The nucleic acids including the polymorphic sites can be  
XX CC used as probes or primers for expressing variant proteins. Detection of  
XX CC the polymorphisms is useful in designing prophylactic and therapeutic  
XX CC regimes customized to underlying abnormalities. The polymorphisms can be  
XX CC used for association studies for hypertension, and in hypertension  
XX CC diagnostic assays. Where the polymorphisms have strong correlation with  
XX CC hypertension, within a gene, they are likely to have a causative role in  
XX CC hypertension. This information can be used to find the precise role of a  
XX CC polymorphism in the disease, and this can be used to identify potential  
XX CC drugs which combat the disease. The polymorphisms can be tested for  
XX CC association with other diseases e.g. agammaglobulinemia, diabetes  
XX CC insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich  
XX CC syndrome, Fabry's disease, familial hypercholesterolemia, polycystic  
XX CC kidney disease, hereditary spherocytosis, von Willebrand's disease,  
XX CC tubercous sclerosis, hereditary hemorrhagica telangiectasia, familial  
XX CC colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and  
XX CC acute intermittent porphyria. The polymorphic forms can also be used in  
XX CC forensics to identify individuals.  
XX SQ Sequence 29 BP; 10 A; 5 C; 5 G; 8 T; 1 other;  
XX  
Alignment Scores:  
Pred. NO.: 299 Length: 29  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 21 Gaps: 0  
US-09-397-967A-17 (1-4) x AAA04425 (1-29)  
QY 1 LysAspTyrTyr 4

Db 7 AAAGACTATAT 18  
RESULT 9  
ABX69473  
ID ABX69473 standard; DNA; 30 BP.  
XX AC ABX69473;  
XX XX  
XX DT 07-MAY-2003 (first entry)  
XX XX Novel Helicobacter pylori gene PCR primer #2444.  
XX DE  
XX KW Protein-protein interaction; ulcer; selected interacting domain;  
XX SID; PCR; primer; ss.  
XX XX Helicobacter pylori.  
XX OS  
XX XX WO200266501-A2.  
XX PN  
XX PD 29-AUG-2002.  
XX XX  
XX PF 28-DEC-2001; 2001WO-BP15428.  
XX PF  
XX PR 02-JAN-2001; 2001US-259302P.  
XX PR  
XX XX (HYBR-) HYBRIGENICS.  
XX PA (INSP) INST PASTEUR.  
XX XX  
XX PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;  
XX PI WPI; 2002-674910/72.  
XX DR  
XX XX  
XX PT New complexes of protein-protein interactions in Helicobacter pylori,  
XX PT useful for identifying modulating compounds for treating or preventing  
XX PT ulcers in mammals -  
XX XX  
XX PS Example 9; Page 564; 642pp; English.  
XX CC The invention describes a complex of protein-protein interactions in  
XX CC Helicobacter pylori selected from 421 complexes given in the  
XX CC specification. The complex of protein-protein interactions are useful  
XX CC for screening for agents which modulate the interaction of proteins.  
XX CC Modulating compounds which binds to a targeted bacterial protein may be  
XX CC used for treating or preventing ulcers in a human or animal. This  
XX CC sequence represents a primer used to isolate polynucleotides encoding  
XX CC Helicobacter pylori proteins for studies on protein-protein  
XX CC interactions.  
XX SQ Sequence 30 BP; 13 A; 4 C; 3 G; 7 T; 3 U; 0 other;  
XX  
Alignment Scores:  
Pred. NO.: 309 Length: 30  
Score: 25.00 Matches: 4  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 24 Gaps: 0  
US-09-397-967A-17 (1-4) x ABX69473 (1-30)  
QY 1 LysAspTyrTyr 4  
Db 17 AAAGATTACTAT 28  
RESULT 10  
AAZ66789/c  
ID AAZ66789 standard; DNA; 47 BP.  
XX AC AAZ66789;  
XX XX  
XX DT 10-SEP-2001 (first entry)  
XX XX

DE Human map-related biallelic marker SEQ ID NO:1136.  
 XX Human genome; biallelic marker; high density disequilibrium map;  
 KW genomic map; haplotype; phenotype; polymorphic base; genotyping;  
 KW haplotyping; hybridisation; identification; characterisation;  
 KW diagnosis; single nucleotide polymorphism; SNP; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT variation replace(24,G)  
 FT /\*tag= a  
 FT /standard\_name= "single nucleotide polymorphism"  
 XX  
 PN WO9954500-A2.  
 XX  
 PD 28-OCT-1999.  
 XX  
 PF 21-APR-1999; 99WO-IB00822.  
 XX  
 PR 21-APR-1998; 98US-0082614.  
 XX  
 PR 23-NOV-1998; 98US-0109732.  
 XX  
 XX (GEST ) GENSET.  
 PA  
 XX  
 PI Cohen D, Blumenfeld M, Chumakov I;  
 XX  
 DR WPI; 2000-013267/01.  
 XX  
 PT Novel biallelic markers used to construct a high density disequilibrium  
 XX map of the human genome -  
 PS Claim 1; Page 473; 2745pp; English.  
 XX  
 CC AA265654 to AA269578 represent human biallelic markers from the present  
 CC invention, which contain a polymorphic base at position 24 of their  
 CC nucleotide sequences. AA269579 to AA277440 represent amplification  
 CC primers for the biallelic markers. The biallelic markers of the  
 CC invention have a variety of uses: they can be used for high density  
 CC mapping of the human genome, and in complex association studies and  
 CC haplotyping studies which are useful in determining the genetic basis  
 CC for disease states. Compositions and methods of the invention can also  
 CC be useful for the identification of the targets for the development of  
 CC pharmaceutical agents and diagnostic methods, as well as the  
 CC characterisation of the differential efficacious responses to and side  
 CC effects from pharmaceutical agents acting on a disease as well as other  
 CC treatment.  
 CC N.B. The SEQ ID NOS 2852, 2913, 2974, 3035, 3096, 3157, 3227, 3297  
 CC and 3367, are not actually given a sequence in the Sequence Listing  
 CC from the present invention.  
 XX  
 SQ Sequence 47 BP; 16 A; 7 C; 7 G; 17 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 478 Length: 47  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 21 Gaps: 0  
 US-09-397-967A-17 (1-4) x AA266789 (1-47)  
 QY 1 LysAspTyrTyr 4  
 DB 47 AAGGATTACTAT 36  
 RESULT 11  
 ID AA175504/c  
 ID AA175504 standard; DNA; 51 BP.  
 XX  
 AC AA175504;  
 XX

DT 09-NOV-2001 (first entry)  
 XX  
 DE Human silent SNP containing nucleic acid SEQ:2445.  
 XX  
 KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
 KW protein therapy; vaccine; probe; diagnostic assay; detection;  
 KW quantitation; restorative therapy; polymorphic; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200140521-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 30-NOV-2000; 2000WO-US32758.  
 XX  
 PR 30-NOV-1999; 99US-0168138.  
 XX  
 PR 29-NOV-2000; 2000US-0726173.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Shinkets RA, Leach M;  
 XX  
 DR WPI; 2001-356160/37.  
 XX  
 PT Polymorphic nucleic acid sequences, useful in genetic testing and  
 PT therapy -  
 XX  
 PS Claim 1; Page 799; 2653pp; English.  
 XX  
 CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide  
 CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
 CC AA173114 to AA175329 represent peptides related to human polymorphic  
 CC polynucleotide sequences. The sequences can be used in gene and protein  
 CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate expression of polymorphic polypeptides.  
 CC For example, (I) may be used to treat disorders by rectifying mutations  
 CC or deletions in a patient's genome that affect the activity of  
 CC polypeptides by expressing inactive proteins or to supplement the  
 CC patients own production of polypeptide. Additionally, (I) and its  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids  
 CC in samples, and therefore which patients may be in need of restorative  
 CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
 CC production of antibodies specific for polymorphic polypeptides. The  
 CC antibodies may also be used to down regulate expression and activity.  
 CC The antibodies may also be used as diagnostic agents for detecting the  
 CC presence of polymorphic polypeptides in samples.  
 XX  
 SQ Sequence 51 BP; 10 A; 11 C; 9 G; 21 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 517 Length: 51  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-397-967A-17 (1-4) x AA175504 (1-51)  
 QY 1 LysAspTyrTyr 4  
 DB 44 AAGGATTATTAC 33  
 RESULT 12  
 ID AA175505/c  
 ID AA175505 standard; DNA; 51 BP.  
 XX  
 AC AA175505;  
 XX  
 DT 09-NOV-2001 (first entry)



XX DE Human silent SNP containing nucleic acid SEQ:2446.  
 XX KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
 KW protein therapy; vaccine; probe; diagnostic assay; detection;  
 KW quantitation; restorative therapy; polymorphic; ds.  
 XX OS Homo sapiens.  
 XX PN WO200140521-A2.  
 XX PD 07-JUN-2001.  
 XX PF 30-NOV-2000; 2000WO-US32758.  
 XX PR 30-NOV-1999; 99US-0168138.  
 XX PR 29-NOV-2000; 2000US-0726173.  
 XX PA (CURA-) CURAGEN CORP.  
 XX XX Shimkets RA, Leach M;  
 XX WPI; 2001-356160/37.  
 XX PT Polymorphic nucleic acid sequences, useful in genetic testing and  
 PT therapy -  
 XX PS Claim 1; Page 800; 2653pp; English.  
 CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide  
 CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
 CC AA53114 to AA53329 represent peptides related to human polymorphic  
 CC polynucleotide sequences. The sequences can be used in gene and protein  
 CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate expression of polymorphic polypeptides.  
 CC For example, (I) may be used to treat disorders by rectifying mutations  
 CC or deletions in a patient's genome that affect the activity of  
 CC polypeptides by expressing inactive proteins or to supplement the  
 CC patients own production of polypeptide. Additionally, (I) and its  
 CC complementary sequences may also be used as DNA probes in diagnostic  
 CC assays to detect and quantitate the presence of similar nucleic acids  
 CC in samples, and therefore which patients may be in need of restorative  
 CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
 CC production of antibodies specific for polymorphic polypeptides. The  
 CC antibodies may also be used to down regulate expression and activity.  
 CC The antibodies may also be used as diagnostic agents for detecting the  
 CC presence of polymorphic polypeptides in samples.  
 XX SQ Sequence 51 BP; 10 A; 12 C; 9 G; 20 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 517 Length: 51  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DE: 22 Gaps: 0  
 US-09-397-967A-17 (1-4) x AA175505 (1-51)  
 QY 1 LysAspTyrTyr 4  
 Db 44 AAGGATTATTAC 33  
 RESULT 13  
 AAC11379/c  
 ID AAC11379 standard; cDNA; 63 BP.  
 XX AC AAC11379;  
 XX AC AAC11379;  
 XX DT 06-OCT-2000 (first entry)  
 XX XX Bos taurus.

DE Human secreted protein 5' EST, SEQ ID NO: 15454.  
 XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 XX OS Homo sapiens.  
 XX PN EP1033401-A2.  
 XX PD 06-SEP-2000.  
 XX PF 21-FEB-2000; 2000EP-0200610.  
 XX PR 26-FEB-1999; 99US-0122487.  
 XX XX (GEST ) GENSET.  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 XX WPI; 2000-500381/45.  
 XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
 XX PS Claim 1; SEQ ID 15454; 71pp + CD-ROM; English.  
 CC The present sequence is one of a large number of 5' ESTs derived from  
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
 CC identified within the present sequence. The 5' ESTs were prepared from  
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
 CC derived from the 5' ends of mRNAs and even in those cases where longer  
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
 CC They are used to obtain upstream regulatory sequences and to design  
 CC expression and secretion vectors.  
 XX SQ Sequence 63 BP; 18 A; 9 C; 8 G; 28 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 635 Length: 63  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DE: 21 Gaps: 0  
 US-09-397-967A-17 (1-4) x AAC11379 (1-63)  
 QY 1 LysAspTyrTyr 4  
 Db 19 AAAGATTACTAC 8  
 RESULT 14  
 AAQ33604  
 ID AAQ33604 standard; DNA; 64 BP.  
 XX AC AAQ33604;  
 XX AC AAQ33604;  
 XX DT 25-MAR-2003 (updated)  
 XX DT 02-FEB-1993 (first entry)  
 XX DE Downstream sequence of microsatellite from clone AGLA296.  
 XX KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;  
 KW genetic mapping; traits; amplification; ss.  
 XX XX Bos taurus.

XX WO9213102-A1.  
 XX PD 06-AUG-1992.  
 XX PF 15-JAN-1992; 92WO-US00340.  
 XX PR 15-JAN-1991; 91US-0642342.  
 XX PA (GENM-) GENMARK.  
 XX PI Georges M, Massey JM;  
 XX PF WPI; 1992-284684/34.  
 XX PT Polymorphic bovine DNA markers - used in genetic identification,  
 XX PT gene mapping, and selective breeding  
 XX PS Table 7; Page 171; 517pp; English.  
 XX  
 CC The sequence is that downstream of a bovine microsatellite sequence  
 CC obtd. by screening a library of bovine MboI DNA fragments of between  
 CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.  
 CC One out of 50 clones cross-hybridised. Assuming independent  
 CC distribution of microsatellites and MboI sites, the frequency of  
 CC (76)n > 9 microsatellites in the bovine genome is estimated at >100,  
 CC 000. The sequence information for ca. 230 such bovine microsatellites  
 CC is summarised in the specification and indexed herein (see below).  
 CC The sequences upstream and downstream of the microsatellite sequence  
 CC were used to generate the required PCR primers for in vitro  
 CC amplification of the corresp. microsatellite (using the program  
 CC OPTIPRIM). The microsatellites may be used to identify individuals,  
 CC for parentage testing, and in the genetic mapping of economic trait  
 CC loci, or genes involved in the determination of economically important  
 CC traits esp. in cattle, to allow selective breeding.  
 CC See also AAQ33501-34437.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 64 BP; 25 A; 10 C; 12 G; 17 T; 0 other;  
 XX

Alignment Scores:  
 Pred. No.: 645 Length: 64  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-397-967A-17 (1-4) x AAQ33604 (1-64)

QY 1 LysAspTyrTyr 4  
 Db 20 AAGGATTATTAC 31

RESULT 15  
 AAF90751/c  
 ID AAF90751 standard; RNA; 71 BP.  
 XX  
 AC AAF90751;  
 XX  
 DT 03-MAY-2001 (first entry)  
 XX  
 DE  
 DE  
 XX  
 KW Human; transforming growth factor beta2; TGFbeta2; SELEX;  
 KW systemic evolution of ligands by exponential enrichment; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200109156-A1.  
 XX  
 PD 08-FEB-2001.  
 XX

PF 26-JUL-2000; 2000WO-US20397.  
 XX  
 PR 29-JUL-1999; 99US-0363939.  
 XX  
 PA (NEXS-) NEXSTAR PHARM INC.  
 XX  
 PI Pagratis N, Lochrie M, Gold L;  
 XX  
 DR WPI; 2001-218217/22.  
 XX  
 XX New RNA ligand to human transforming growth factor beta2, useful as  
 XX pharmaceuticals, diagnostics and as immunochemical reagents -  
 XX Claim 1; Page 61; 178pp; English.  
 XX  
 CC The present invention relates to non-naturally occurring, high-affinity  
 CC RNA ligands to human transforming growth factor beta2 (TGFbeta2). The  
 CC oligonucleotide ligands were identified by the SELEX method (SELEX stands  
 CC for Systemic Evolution of Ligands by Exponential Enrichment). The  
 CC oligonucleotide ligands are useful in any process in which binding to  
 CC TGFbeta2 is required. The ligands may be useful as pharmaceuticals,  
 CC diagnostics, imaging agents and immunochemical reagents. The present  
 CC sequence is an oligonucleotide used in the present invention.  
 XX  
 SQ Sequence 71 BP; 12 A; 22 C; 16 G; 21 U; 0 other;  
 XX

Alignment Scores:  
 Pred. No.: 713 Length: 71  
 Score: 25.00 Matches: 4  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0

US-09-397-967A-17 (1-4) x AAF90751 (1-71)

QY 1 LysAspTyrTyr 4  
 Db 26 AAGGATTATTAC 15

Search completed: February 4, 2004, 03:08:18  
 Job time : 3.90481 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 13:04:22 ; Search time 873 Seconds

(without alignments)  
4637.242 Million cell updates/sec

Title: US-09-397-967A-16

Perfect score: 1099

Sequence: 1 MAPSEETLPQPSCLSS.....RPAPATLSQLDPLWRGPG 1099

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4891994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPto\_spool/US09397967/runat\_03022004\_175718\_28991/app\_query.fasta\_1.1287  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=olip2n.rnpb -MINMATCH=0.1  
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEADSIZE=500 -MINLEN=1  
-MAXLEN=200000000 -USER=US09397967 @CGN 1 1 331 @runat\_03022004\_175718\_28991  
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/FCI\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
------------	-------------	--------	----	-------------

1	62	5.6	3620	12	US-10-305-720-1055	Sequence 1055, Ap
2	62	5.6	5118	13	US-09-814-353-19356	Sequence 19356, A
3	58	5.3	203	9	US-09-728-446-1413	Sequence 1413, Ap
4	46	4.2	439	11	US-09-918-995-11571	Sequence 11571, A
5	28	2.5	438	11	US-09-918-995-10879	Sequence 10879, A
6	21	1.9	65	13	US-09-908-975-30425	Sequence 30425, A
7	20	1.8	860	13	US-10-027-632-8229	Sequence 8229, Ap
8	20	1.8	860	14	US-10-027-632-8229	Sequence 8229, Ap
9	18	1.6	364	10	US-09-771-161A-17	Sequence 17, Appl
10	16	1.5	3731	12	US-10-191-803-200	Sequence 200, App
11	16	1.5	4078	12	US-10-305-720-297	Sequence 297, App
12	16	1.5	7126	12	US-10-062-674-2214	Sequence 2214, Ap
13	12	1.1	151	10	US-09-982-610-11	Sequence 11, Appl
14	12	1.1	215	13	US-10-029-386-26806	Sequence 26806, A
15	12	1.1	294	9	US-09-864-761-21086	Sequence 21086, A
16	12	1.1	367	9	US-09-864-761-4335	Sequence 4335, Ap
17	12	1.1	372	10	US-09-878-574-1581	Sequence 1581, Ap
18	12	1.1	477	11	US-09-918-995-2066	Sequence 2066, Ap
19	12	1.1	559	15	US-10-066-543-1870	Sequence 1870, Ap
20	12	1.1	562	13	US-10-029-386-13106	Sequence 13106, A
21	12	1.1	637	13	US-10-027-632-247078	Sequence 247078, A
22	12	1.1	637	14	US-10-027-632-247078	Sequence 247078, A
23	12	1.1	917	11	US-09-991-936-1577	Sequence 1577, Ap
24	12	1.1	927	13	US-10-430-797-3	Sequence 3, Appl
25	12	1.1	1509	13	US-10-205-194-10	Sequence 10, Appl
26	12	1.1	1583	11	US-09-971-708-6	Sequence 6, Appl
27	12	1.1	1713	13	US-10-187-900-1	Sequence 1, Appl
28	12	1.1	1987	13	US-10-100-217-1	Sequence 1, Appl
29	12	1.1	1989	15	US-10-084-817-341	Sequence 341, App
30	12	1.1	2000	9	US-09-977-269-1	Sequence 1, Appl
31	12	1.1	2000	10	US-09-977-260-1	Sequence 1, Appl
32	12	1.1	2000	11	US-09-977-261-1	Sequence 1, Appl
33	12	1.1	2065	13	US-10-103-380A-1	Sequence 1, Appl
34	12	1.1	2829	13	US-10-430-797-1	Sequence 1, Appl
35	12	1.1	2982	13	US-10-187-958-2	Sequence 2, Appl
36	12	1.1	3107	10	US-09-954-456-1128	Sequence 1128, Ap
37	12	1.1	3136	13	US-10-412-277-1	Sequence 1, Appl
38	12	1.1	3170	13	US-09-971-392-66	Sequence 66, Appl
39	12	1.1	3220	13	US-10-430-797-5	Sequence 5, Appl
40	12	1.1	3348	10	US-09-982-610-34	Sequence 34, Appl
41	12	1.1	3370	10	US-09-967-768A-144	Sequence 144, App
42	12	1.1	3370	13	US-10-354-358-101	Sequence 101, App
43	12	1.1	3370	13	US-10-210-120-19	Sequence 19, Appl
44	12	1.1	3393	9	US-09-922-138-3	Sequence 3, Appl
45	12	1.1	3583	13	US-10-237-496-103	Sequence 103, App

ALIGNMENTS

RESULT 1  
US-10-305-720-1055  
; Sequence 1055, Application US/10305720  
; Publication No. US20040010136A1  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.  
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression  
; FILE REFERENCE: PA-0002-1 CON  
; CURRENT APPLICATION NUMBER: US/10/305,720  
; CURRENT FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 09/016,434  
; PRIOR FILING DATE: 1998-01-30  
; NUMBER OF SEQ ID NOS: 1490  
; SOFTWARE: PERL Program  
; SEQ ID NO 1055  
; LENGTH: 3620  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g1039418  
US-10-305-720-1055

Alignment Scores:

Pred. No.: 9,75e-52 Length: 3620  
 Score: 62.00 Matches: 96  
 Percent Similarity: 97.96% Conservatives: 0  
 Best Local Similarity: 97.96% Mismatches: 1  
 Query Match: 5.64% Indels: 2  
 DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-305-720-1055 (1-3620)

QY 938 ArgArgCysValHisArgAspLeuAlaAArgAenilleLeuValGluSerGluAlaHis 957  
 DB 2922 CGCCGCTGCTGACCGGACCTGCGCCGCGAATCATCTCTGAGAGCGAGGACAC 2981  
 QY 958 VallylsilleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977  
 DB 2982 GTCAGATCGCTGACTTCGCGCTAGCTAGCTGCGCCGCTGAC-AAAGACTACTACGT 3040  
 QY 977 lValArgGluProGlyGlnSerProillePheTrpTyrAlaProGluSerLeuSerAspAs 997  
 DB 3041 GGTCCGCGAGCCAGCCAGCCCATTTCTGGTATGCCCCGGAATCCCTCTCGGACAA 3100  
 QY 997 nilePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017  
 DB 3101 CATCTTCTCTCGCCAGTCAGACGTCTGGAGCTTCGGGGTCTGCTGTACGAGCTCTTCAC 3160  
 QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034  
 DB 3161 CTACTGCGACAAAGCTGAGCCCTCGCGCGAGTTCTTGGGATGATGGGA 3212

RESULT 2

US-09-814-353-19356

; Sequence 19356, Application US/09814353  
 ; Publication No. US20030165831A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Thompson, Pamela  
 ; APPLICANT: Lillie, James  
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
 ; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 ; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
 ; FILE REFERENCE: MRI-006B  
 ; CURRENT APPLICATION NUMBER: US/09/814,353  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/191,031  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: US 60/207,124  
 ; PRIOR FILING DATE: 2000-05-25  
 ; PRIOR APPLICATION NUMBER: US 60/211,940  
 ; PRIOR FILING DATE: 2000-06-15  
 ; PRIOR APPLICATION NUMBER: US 60/216,820  
 ; PRIOR FILING DATE: 2000-07-07  
 ; PRIOR APPLICATION NUMBER: US 60/220,661  
 ; PRIOR FILING DATE: 2000-07-25  
 ; PRIOR APPLICATION NUMBER: US 60/257,672  
 ; PRIOR FILING DATE: 2000-12-21  
 ; NUMBER OF SEQ ID NOS: 22037  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 19356  
 ; LENGTH: 5118  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: 1, 2, 3, 4, 5114, 5115, 5116, 5117, 5118  
 ; OTHER INFORMATION: n = A,T,C or G

US-09-814-353-19356

Alignment Scores:  
 Pred. No.: 1.33e-51 Length: 5118  
 Score: 62.00 Matches: 96  
 Percent Similarity: 97.96% Conservatives: 0  
 Best Local Similarity: 97.96% Mismatches: 1  
 Query Match: 5.64% Indels: 2

DB: 13 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-814-353-19356 (1-5118)

QY 938 ArgArgCysValHisArgAspLeuAlaAArgAenilleLeuValGluSerGluAlaHis 957  
 DB 2923 CGCCGCTGCTGACCGGACCTGCGCCGCGAATCATCTCTGAGAGCGAGGACAC 2982  
 QY 958 VallylsilleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977  
 DB 2983 GTCAGATCGCTGACTTCGCGCTAGCTAGCTGCGCCGCTGAC-AAAGACTACTACGT 3041  
 QY 977 lValArgGluProGlyGlnSerProillePheTrpTyrAlaProGluSerLeuSerAspAs 997  
 DB 3042 GGTCCGCGAGCCAGCCAGCCCATTTCTGGTATGCCCCGGAATCCCTCTCGGACAA 3101  
 QY 997 nilePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017  
 DB 3102 CATCTTCTCTCGCCAGTCAGACGTCTGGAGCTTCGGGGTCTGCTGTACGAGCTCTTCAC 3161  
 QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034  
 DB 3162 CTACTGCGACAAAGCTGAGCCCTCGCGCGAGTTCTTGGGATGATGGGA 3213

RESULT 3

US-09-728-446-1413

; Sequence 1413, Application US/09728446  
 ; Patent No. US20020081668A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Friedrich, Glenn  
 ; APPLICANT: Zambrowicz, Brian  
 ; APPLICANT: Sands, Arthur T.  
 ; TITLE OF INVENTION: No. US20020081668A1 Murine Polynucleotide Sequences  
 ; TITLE OF INVENTION: and Mutant Cells and Mutant Animals Defined Thereby  
 ; FILE REFERENCE: LEX-0101-USA  
 ; CURRENT APPLICATION NUMBER: US/09/728,446  
 ; CURRENT FILING DATE: 2000-11-30  
 ; PRIOR APPLICATION NUMBER: US 60/168,270  
 ; PRIOR FILING DATE: 1999-12-01  
 ; NUMBER OF SEQ ID NOS: 1461  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1413  
 ; LENGTH: 203  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 ; US-09-728-446-1413

Alignment Scores:  
 Pred. No.: 8.04e-49 Length: 203  
 Score: 58.00 Matches: 58  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.28% Indels: 0  
 DB: 9 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-728-446-1413 (1-203)

QY 1 MetAlaProProSerGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20  
 DB 18 ATGGCACCTCCAAAGTGGAGAGACCTCTGATCCCTCAGCGCTTTCGACGCTCTATCC 77  
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
 DB 78 TCAGAGCGAGGAGCCCTGCGATGTCTCTCTCCCGGGGACCTGGCGCTCCCGACGGA 137  
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAa 58  
 DB 138 TTGTCTATCTCTTTTGGGGACTACTTGGCTGAGATTATGTGTGCGAGCTGCC 191

RESULT 4

US-09-918-995-11571

; Sequence 11571, Application US/09918995  
 ; Publication No. US20030073623A1

; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 ; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
 ; FILE REFERENCE: 20411-756  
 ; CURRENT APPLICATION NUMBER: US/09/918,995  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: US/09/235,076  
 ; PRIOR FILING DATE: 1999-01-20  
 ; NUMBER OF SEQ ID NOS: 38054  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 11571  
 ; LENGTH: 439  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(439)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-918-995-11571

Alignment Scores:  
 Pred. No.: 1.92e-36 Length: 439  
 Score: 46.00 Matches: 46  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.19% Indels: 0  
 DB: 11 Gaps: 0

US-09-397-967a-16 (1-1099) x US-09-918-995-11571 (1-439)

Qy 989 TyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSerPhe 1008  
 Db 39 TATGCCCGGATCCCTCGGACACACATCTCTCTGCCAGTCAGAGCTTGAGCTTC 98  
 Qy 1009 GlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGlu 1028  
 Db 99 GGGGTGCTGCTGTACGAGCTCTTACCTACTACGACAAAAGCTCAGGCCCTCGGCCGAG 158  
 Qy 1029 PheLeuArgMetMetGly 1034  
 Db 159 TTCCTGCGGATGATGGGA 176

RESULT 5  
 US-09-918-995-10879  
 ; Sequence 10879, Application US/09918995  
 ; Publication No. US20030073623A1

; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc.  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
 ; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
 ; FILE REFERENCE: 20411-756  
 ; CURRENT APPLICATION NUMBER: US/09/918,995  
 ; CURRENT FILING DATE: 2001-07-30  
 ; PRIOR APPLICATION NUMBER: US/09/235,076  
 ; PRIOR FILING DATE: 1999-01-20  
 ; NUMBER OF SEQ ID NOS: 38054  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 10879  
 ; LENGTH: 438  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: (1)...(438)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-918-995-10879

Alignment Scores:  
 Pred. No.: 2.54e-18 Length: 438  
 Score: 28.00 Matches: 28  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 2.55% Indels: 0  
 DB: 11 Gaps: 0

US-09-397-967a-16 (1-1099) x US-09-918-995-10879 (1-438)

Qy 1007 SerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSer 1026  
 Db 96 AGCTTCGGGGTCTGCTGTACGAGCTCTTACCTACTACGACAAAAGCTCAGGCCCTCG 155  
 Qy 1027 AlaGluPheLeuArgMetMetGly 1034  
 Db 156 GCCGAGTTCCTGCGGATGATGGGA 179

RESULT 6

US-09-908-975-30425  
 ; Sequence 30425, Application US/09908975  
 ; Publication No. US20030165843A1

; GENERAL INFORMATION:  
 ; APPLICANT: SHOSHAN, Avi  
 ; APPLICANT: WASSERMAN, Alon  
 ; APPLICANT: MINTZ, Eli  
 ; APPLICANT: MINTZ, Laet  
 ; APPLICANT: FAIGLER, Simchon  
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
 ; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME  
 ; FILE REFERENCE: 36688-0005  
 ; CURRENT APPLICATION NUMBER: US/09/908,975  
 ; CURRENT FILING DATE: 2001-07-20  
 ; PRIOR APPLICATION NUMBER: US 60/287,724  
 ; PRIOR FILING DATE: 2001-05-02  
 ; PRIOR FILING DATE: 2000-07-28  
 ; NUMBER OF SEQ ID NOS: 32337  
 ; SOFTWARE: Patent in version 3.0  
 ; SEQ ID NO 30425  
 ; LENGTH: 65  
 ; TYPE: DNA  
 ; ORGANISM: Mus musculus  
 US-09-908-975-30425

Alignment Scores:  
 Pred. No.: 5.2e-12 Length: 65  
 Score: 21.00 Matches: 21  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 1.91% Indels: 0  
 DB: 13 Gaps: 0

US-09-397-967a-16 (1-1099) x US-09-908-975-30425 (1-65)

Qy 739 LeuProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAsp 758  
 Db 3 CTGCCCGCTCTCAAAATGACAGAACTGCGGGAGCTTATCACAGTCATGCGCGTATGAT 62  
 Qy 759 Pro 759  
 Db 63 CCT 65

RESULT 7

US-10-027-632-8229/c  
 ; Sequence 8229, Application US/10027632  
 ; Publication No. US20030204075A9

; GENERAL INFORMATION:  
 ; APPLICANT: Wang, David G.  
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
 ; TITLE OF INVENTION: Polymorphisms in the Human Genome  
 ; FILE REFERENCE: 108827.129  
 ; CURRENT APPLICATION NUMBER: US/10/027,632  
 ; CURRENT FILING DATE: 2002-04-30  
 ; PRIOR APPLICATION NUMBER: US 60/218,006  
 ; PRIOR FILING DATE: 2000-07-12  
 ; PRIOR APPLICATION NUMBER: US 60/198,676  
 ; PRIOR FILING DATE: 2000-04-20

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8229
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(860)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-8229

Alignment Scores:
Pred. No.: 5.23e-10 Length: 860
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.82% Indels: 0
DB: 13 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-027-632-8229 (1-860)
QY 889 ProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp 908
Db 366 CCAGGCCCGCAGAGCCTGGCGTGGTCATGGAGTACTGCCCGCAGCGCTGCTTGGCGGAC 307

RESULT 8
US-10-027-632-8229/c
; Sequence 8229, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8229
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(860)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-8229

Alignment Scores:

```

```

Pred. No.: 5.23e-10 Length: 860
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.82% Indels: 0
DB: 14 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-027-632-8229 (1-860)
QY 889 ProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp 908
Db 366 CCAGGCCCGCAGAGCCTGGCGTGGTCATGGAGTACTGCCCGCAGCGCTGCTTGGCGGAC 307

RESULT 9
US-09-771-161A-17
; Sequence 17, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 17
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-771-161A-17

Alignment Scores:
Pred. No.: 2.51e-08 Length: 364
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.64% Indels: 0
DB: 10 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-771-161A-17 (1-364)
QY 1 MetAlaProSerGluThrProLeuIleProGlnArgSerCysSerLeu 18
Db 96 ATGCACCTCCAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTTCATGCAGCCTC 149

RESULT 10
US-10-191-803-200
; Sequence 200, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5090US
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611

```

FILE REFERENCE: 2002-05-06  
CURRENT APPLICATION NUMBER: US/10/062,674  
PRIOR FILING DATE: 2002-01-30  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 200  
LENGTH: 3731  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM\_031514  
US-10-191-803-200

Alignment Scores:  
Pred. No.: 2,05e-05 Length: 3731  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-191-803-200 (1-3731)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
DB 3259 TCAGATGTTGGAGCTTTGGTGGTTCGTACGAACATTTTCACGTAT 3306

## RESULT 11

US-10-305-720-297

Sequence 297, Application US/10305720

Publication No. US20040010136A1

GENERAL INFORMATION:

APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.

TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression

FILE REFERENCE: PA-0002-1 CON

CURRENT APPLICATION NUMBER: US/10/305,720

CURRENT FILING DATE: 2002-11-26

PRIOR APPLICATION NUMBER: 09/016,434

PRIOR FILING DATE: 1998-01-30

NUMBER OF SEQ ID NOS: 1490

SOFTWARE: PERL Program

SEQ ID NO 297

LENGTH: 4078

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20040010136A1 179527

FEATURE:

NAME/KEY: unsure

LOCATION: (1) ... (4078)

OTHER INFORMATION: a, t, c, g, or other

US-10-305-720-297

Alignment Scores:  
Pred. No.: 2,22e-05 Length: 4078  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-305-720-297 (1-4078)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
DB 3441 TCAGATGTTGGAGCTTTGGTGGTTCGTATGAACATTTTCACATAC 3488

## RESULT 12

US-10-062-674-2214

Sequence 2214, Application US/10062674

Publication No. US20040005559A1

GENERAL INFORMATION:

APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.

TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS

FILE REFERENCE: PA-0026-1 CIP  
CURRENT APPLICATION NUMBER: US/10/062,674  
PRIOR FILING DATE: 2002-01-30  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 200  
LENGTH: 3731  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20040005559A1 903485.2  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1) ... (7126)  
OTHER INFORMATION: a, t, c, g, or other  
US-10-062-674-2214

Alignment Scores:  
Pred. No.: 3,64e-05 Length: 7126  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 12 Gaps: 0

US-09-397-967A-16 (1-1099) x US-10-062-674-2214 (1-7126)

QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
DB 3606 TCAGATGTTGGAGCTTTGGTGGTTCGTATGAACATTTTCACATAC 3653

## RESULT 13

US-09-982-610-11

Sequence 11, Application US/09982610

Patent No. US20020146420A1

GENERAL INFORMATION:

APPLICANT: Genentech, Inc.

Bennett, Brian D.

Goeddel, David

Lee, James M.

Matthews, William

Tsai, Siao Ping

Wood, William I.

TITLE OF INVENTION: PROTEIN TYROSINE KINASE AGONIST ANTIBODIES

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/982,610

FILING DATE: 17-Oct-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/446,648

FILING DATE: 1996-MAY-23

APPLICATION NUMBER: 08/222616

FILING DATE: 04-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Lee, Wendy M.

REGISTRATION NUMBER: 40,378

REFERENCE/DOCKET NUMBER: P0821P3PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:  
LENGTH: 151 base pairs  
TYPE: Nucleic Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
US-09-982-610-11

Alignment Scores:  
Pred. No.: 0.0126 Length: 151  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 10 Gaps: 0

US-09-397-967a-16 (1-1099) x US-09-982-610-11 (1-151)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 1 GTGCACAGGATCTCGGGCTCGGAACATCTCGTC 36

RESULT 14

US-10-029-386-26806  
; Sequence 26806, Application US/10029386  
; Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharron G.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 26806  
LENGTH: 215  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR3.3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.69  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.58  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.35  
OTHER INFORMATION: SWISSPROT HIT: Q07494, EVALUE 2.00e-35  
OTHER INFORMATION: NT HIT: AF037331.1, EVALUE 1.00e-117  
OTHER INFORMATION: EST\_HUMAN HIT: BG481991.1, EVALUE 2.00e-39

US-10-029-386-26806

Alignment Scores:  
Pred. No.: 0.0173 Length: 215  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.09% Indels: 0  
DB: 13 Gaps: 0

US-09-397-967a-16 (1-1099) x US-10-029-386-26806 (1-215)

QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuVal 952

Db 88 GTGCATCGGACCTCGGCTCGTAGGAACATCTCGTC 123

RESULT 15

US-09-864-761-21086  
; Sequence 21086, Application US/09864761

Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Hanzel, David K.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,587  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 21086  
LENGTH: 294  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC010899.2  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.6  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.7  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3  
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 3.8  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.3  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4  
OTHER INFORMATION: SWISSPROT HIT: Q91694, EVALUE 2.00e-35  
OTHER INFORMATION: NT HIT: GI4758279, EVALUE 1.00e-115  
OTHER INFORMATION: EST\_HUMAN HIT: AW965297.1, EVALUE 2.00e-26

Alignment Scores:  
Pred. No.: 0.0228 Length: 294  
Score: 12.00 Matches: 12







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 00:50:11 ; Search time 11745.3 Seconds  
(without alignments)  
3827.903 Million cell updates/sec

Title: US-09-397-967A-16

Perfect score: 5860

Sequence: 1 MAPSEETPLIPQSCSLSS.....RPAFATLSPQLDPLWRGRPG 1099

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09397967/runat\_03022004\_175631\_28576/app\_query.fasta\_1.1486  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OJFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967@cgn\_1\_17598@runat\_03022004\_175631\_28576 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hcg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5854	99.9	3723	10	MUSPTYKIN	L32955 Mouse prote
2	5687.5	97.1	3528	10	MUSJAK3A	L40172 Mus musculu
3	5380.5	91.8	3778	10	RATJAK3	D28508 Rattus norv
4	5223	89.1	4016	10	MUSJAK3H	L33788 Mus musculu
5	4684	79.9	4064	6	AX203081	AX203081 Sequence
6	4684	79.9	4064	9	HSU09607	U09607 Human JAK f
7	4597	78.4	3620	6	AR270492	AR270492 Sequence
8	4597	78.4	3620	9	HSU31601	U31601 Human tyros
9	4091.5	69.8	218208	2	AC073700	AC073700 Mus muscu
10	4091.5	69.8	223734	2	AC073750	AC073750 Mus muscu
11	3935	67.2	10326	10	MMU71201	U71201 Mus musculu
12	3692	63.0	216847	2	AC103459	AC103459 Rattus no
13	3336	56.9	3644	5	AF034576	AF034576 Gallus ga
14	3288.5	55.1	2914	6	AX203084	AX203084 Sequence
15	2713	45.3	13562	9	HSU70065	U70065 Human JAK3
16	2637.5	45.0	4812	4	AB006011	AB006011 Sus scrof
17	2636.5	45.0	3731	10	RNU13396	U13396 Rattus norv
18	2629	44.9	3399	6	AX695805	AX695805 Sequence
19	2629	44.9	3435	6	ARI29823	ARI29823 Sequence
20	2629	44.9	4161	9	AF005216	AF005216 Homo sapi
21	2629	44.9	4216	9	BC039695	BC039695 Homo sapi
22	2629	44.9	5097	6	AX695804	AX695804 Sequence
23	2628	44.8	3500	9	AF001362	AF001362 Homo sapi
24	2621	44.7	5117	6	AX354568	AX354568 Sequence
25	2621	44.7	5117	9	AF058925	AF058925 Homo sapi
26	2619.5	44.7	3629	6	ARI36148	ARI36148 Sequence
27	2619.5	44.7	3629	6	ARI43945	ARI43945 Sequence
28	2619.5	44.7	3629	6	I92575	I92575 Sequence 8
29	2613.5	44.6	3390	6	AX695802	AX695802 Sequence
30	2613.5	44.6	3629	6	AX695801	AX695801 Sequence
31	2613.5	44.6	3629	10	MUSPTK	L16956 Mouse prote
32	2606.5	44.5	4387	5	DRE5690	AJ005690 Danio rer
33	2594	44.3	4078	6	AR269734	AR269734 Sequence
34	2523.5	43.1	3739	5	AF148993	AF148993 Cyprinus
35	2511	42.8	4042	4	AB036337	AB036337 Sus scrof
36	2460.5	42.0	2881	9	BC028068	BC028068 Homo sapi
37	2382.5	40.7	41006	9	AC007201	AC007201 Homo sapi
38	2360	40.3	3495	6	AR048200	AR048200 Sequence
39	2360	40.3	3495	6	AR068116	AR068116 Sequence
40	2360	40.3	3495	6	AR071176	AR071176 Sequence
41	2360	40.3	3495	6	ARI77639	ARI77639 Sequence
42	2360	40.3	3495	6	I62349	I62349 Sequence 2
43	2360	40.3	3495	6	I87981	I87981 Sequence 2
44	2350	40.1	4599	4	AB036336	AB036336 Sus scrof
45	2339	39.9	21323	9	AF513860	AF513860 Homo sapi

# ALIGNMENTS

RESULT 1

MUSPTYKIN 3723 bp mRNA linear ROD 21-SEP-1994  
 LOCUS Mouse protein tyrosine kinase (jak3) mRNA, complete cds.  
 DEFINITION  
 ACCESSION L32955.1 GI:529238  
 VERSION interleukin 2; interleukin 4; Janus kinase; protein tyrosine kinase.  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 3723)  
 AUTHORS Withuhn, B.A., Silvennoinen, O., Miura, O., Lai, K.S., Cwik, C., Liu, E.T. and Inle, J.N.  
 TITLE Involvement of the Jak-3 Janus kinase in signalling by interleukins 2 and 4 in lymphoid and myeloid cells  
 JOURNAL Nature 370 (6485), 153-157 (1994)  
 MEDLINE 94294024  
 PUBMED 8022486  
 COMMENT Original source text: Mus musculus (strain BALB/c, sub\_species domesticus) cDNA to mRNA.  
 This entry has been reported under the accession number L329555 in the Nature article.  
 FEATURES  
 Location/Qualifiers  
 1..3723  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /sub\_species="domesticus"  
 /db\_xref="taxon:10090"  
 /cell\_line="Abelson transformed pre-B-cell line 22D6"  
 /cell\_type="lymphoid and myeloid"  
 1..3723  
 /gene="jak3"  
 1..12  
 /genes="jak3"  
 13..3312  
 /genes="jak3"  
 /standard\_name="Janus kinase, mujak3"  
 /codon\_start=1  
 /evidence=experimental  
 /product="protein tyrosine kinase"  
 /protein\_id="AA021565.1"  
 /db\_xref="GI:529239"  
 /translations="MAPPSETPLIPORCSLSSBAGALHVLPPRGPPORLSFS  
 FGYLADLICVRAKACAILPVYHSLPALATEDPFCWPPSPHFICIEDVDTQVLVYRL  
 RFYPPWFGLETCHRFGLKDLTSAIDLHLVSHLPAQRSDILVSGRLPVGLSWKEQG  
 EFUSLAVLDLQMARQAQRPGLLTVSYKACLPSLRDVIQGVNFRVRRIRRTVW  
 LALLPCGRLPFRYALMAKYILDLERLHPAATETTFVGLPGAQEEFGLAEFPGLPEA  
 PWSNDELFTQDFPIVDVSIQAPRVPGAGEHLVTVTRMDGHEFLAEFPGLPEA  
 LSFVALVDGFRILICDSRHVFCKEVAPPRLLEEDVCHGPITLDFAIHLKKAAGSLP  
 GTYLRSPPQDYSFLTACTVQTPGLPDYKGLIRQDPGAFSLVGLSQPHRSRLRELL  
 AACWNSLRVDGAALYLTCCAPRPKEKSNLIVRRGNPAPAPGCPSCCALUTQSF  
 HTPTPSLEHNLGHGSFTKIPGRSERREVDGETHDSVLLKVPMSRRHNCMESFLE  
 AASLMQVSYPHLVILGVCWAGDSIMVQEFVYLGAIDMYLRKRLHLSWAKLQVTK  
 OLAVALNLDKGLPHGNVSARKVLLAREGDNPPFIKLSDPGVSTVLSLEMLTDR  
 IPWVAPCEAOATCLLEADKQFGATWVEVFORGAHITSLRPAKKLKEYEDQGLP  
 ALKWTLAGLITQCMYDPPRRSPFRAILDLNGLITSYDELLSDTPTGPIPSRDLIC  
 VAGAQYACDDPAIFERHLKYISLIGKNGFSGVELCDPLGNDTPIVAVKQLQHS  
 VPQQQDFQREIQLKALHSDFIVKRGVSYGFRSLRMEYLPSCGLRDLQLQRH  
 GLTDRLLFAWQITCKMEYLGARRCVHRDLAARNILVSEAHVKIADFGLAKLLPLG  
 KDYVVRPEQSPFIWYAPESLSDNIPFSROSDVMSFGVVLVYELFTYCDKSCSPSAPFL  
 RWMGPEREGPPLCRILLEAEGRRLLPPPTCPTEVQELMQLCWAPEPHDRPAPATLSLP  
 QLDFLRGRPG"  
 13..3309  
 /gene="jak3"  
 /product="protein tyrosine kinase"  
 /standard\_name="Janus kinase, mujak3"  
 /evidence=experimental  
 3313..3723  
 /gene="jak3"  
 3723  
 /gene="jak3"  
 mat\_peptide  
 3'UTR  
 polyA\_site

BASE COUNT 711 a 1162 c 1079 g 771 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 3723  
 Score: 5854.00 Matches: 1098  
 Percent Similarity: 99.91% Conservative: 0  
 Best Local Similarity: 99.91% Mismatches: 1  
 Query Match: 99.90% Indels: 0  
 DB: 10 Gaps: 0

US-09-397-967A-16 (1-1099) x MUSPTYKIN (1-3723)

QY 1 MetAlaProProSerGluGluThrProLeuLeuLeuProGlnArgSerCysSerLeuSerSer 20  
 |||||  
 Db 13 ATGGCACTCCCAAGTGAGGAGACACCTCTGATCCCTCAGCGCTCTTGACGCTCTCATCC 72  
 |||||  
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
 |||||  
 Db 73 TCAGAGGAGGAGGAGCTTGATGTGCTCTCTCCCGGGGAGCTGGCGCTCCCGGCGA 132  
 |||||  
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValargalaAlaAla 60  
 |||||  
 Db 133 TTGTCACTCTCTTTTGGGAGCTACTTTGGCTGAGGATTTATGTGTGGAGCTGCCAAGGCC 192  
 |||||  
 QY 61 CysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 |||||  
 Db 193 TGTGCTATCTGCTGCTTTATCATTCGCTTTTCGCTTGCCACTGAGGACTTCTCTTGC 252  
 |||||  
 QY 81 TrpPheProProSerHisPheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 |||||  
 Db 253 TGGTTTCCCAAGCACATCTTCTCATAGAGGAGCTGGACACTCAAGTCTTGGTCTAC 312  
 |||||  
 QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
 |||||  
 Db 313 AGGCTACGCTTTTATTTCCCTGACTGTTTGGGCTGGAGACATGTCAACGCTTTGGGCTG 372  
 |||||  
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
 |||||  
 Db 373 CGCAAGATTTGACAGTGCCATCCTTGACTTACATGTTTAGACATCTCTTTGCTCAG 432  
 |||||  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
 |||||  
 Db 433 CACCGCAGTGACCTGGTGTAGTGGCGCTCCCGTGGGCTTAGCATGAAGGAGCAGGGA 492  
 |||||  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGluMetAlaArgGluGlnAlaGlnArg 180  
 |||||  
 Db 493 GAGTTCTGAGCCTGGCGCTGGAGCTTGAGCTTGAGCTGCTGAGCAGGCCAGCGC 552  
 |||||  
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
 |||||  
 Db 553 CCAGGAGAGCTGCTGAAGAGGTCAGTTACAAAGCTGTCTGCCCGCCAGCTGGCGCAT 612  
 |||||  
 QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
 |||||  
 Db 613 GTGATCCAGGGCCAGAACTTCGTGACACGACGAGCGCATCCGAGGACCGGTGGTGGCG 672  
 |||||  
 QY 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240  
 |||||  
 Db 673 CTGCTCGGTGTGGTGGCTTGCCGAGCGCCAGCGCTCATGGCCAAGTATATTTCTG 732  
 |||||  
 QY 241 AspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260  
 |||||  
 Db 733 GACCTGGAGGGCTACATCCAGCGGCCACCCAGAGACCTTTCGCTGTGGGGCTCCCGGGC 792  
 |||||  
 QY 261 AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSer 280  
 |||||  
 Db 793 GCCCAGGAGGAGCGGGGCTTCTGCGTGTGGGAGGAGCAACCGCATCTCCCTTGAGTCTG 852  
 |||||  
 QY 281 AsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300  
 |||||  
 Db 853 AACGACAGCTTTTCCAGACCTTCTGTGACTTTCCGGAATCGTGGATGTGAGCATCAAT 912  
 |||||

QY 301 GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAsp 320  
DB 913 CAGGCCCCACAGTGTGGTCCGCGAGGGAGCACCGCGTCTACTGTCAACAGGATGGAC 972  
QY 321 GlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu 340  
DB 973 GGCCACATCCTTGGAAAGCGAGTTCTCGGGGCTGCGCTGAGGCGCTGTCTTTCGTGGCGCTC 1032  
QY 341 ValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysIleGluValAla 360  
DB 1033 GTGGATGGGTACTTCGCGCTGATCTGCGACTCCAGGCATTATTTCTGCAAGGAGGTGGCG 1092  
QY 361 ProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPhe 380  
DB 1093 CCGCCAAAGGCTGCTGGAGAGGAGCGGACGTGTGCCATGGACCCATCAGCTTAGACTTT 1152  
QY 381 AlalIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400  
DB 1153 GCCATFCCAAGCTGAAGGCCGCTGGCTCCCTCCAGGACCATATATTCTCCGCGCGCAGC 1212  
QY 401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420  
DB 1213 CCGCAGGACTATGACAGCTTTCTTTACCGCTCGCTCCAGACTCTCTTGGCCCCGAC 1272  
QY 421 TyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440  
DB 1273 TACAAGGGCTGCCTCATCCGCCAGGACCCACCGGGGCTTTCTCCCTGGTTGGCCTCAGC 1332  
QY 441 GlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrPheAsnSerGlyLeuArgVal 460  
DB 1333 CAGCCCCACAGAAAGCTCGGGAGCTGCTTGCGAGCTGCTGGAAATTTCTGGGCTGGAGTA 1392  
QY 461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsn 480  
DB 1393 GACGGTGTGCTGCTGTACTTAACATCTCTGCTCGCTCCAGACCCCAAGAAAGTCCCAAT 1452  
QY 481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500  
DB 1453 TTGATCGTGGTGCGAAGGGGCTGCAACCCCGCGCTGCCCCCTGGCTGCTCCCGTCTGTC 1512  
QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluThrHisGlu 520  
DB 1513 TGTGGCGCTGACACAGCTGAGCTTCCACACAATTCACAACGGAAGCTGCGAGTGGCAGCAG 1572  
QY 521 AsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp 540  
DB 1573 AACCTGGGTACCGTTCTTTTACCAAGATCTTCCGTGGCAGCAGCGGAGGTCGTGGAT 1632  
QY 541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560  
DB 1633 GGTGAGACACATGACTCGGAAGTCTCTCTGAAGGTATGGACTCCAGACATCGGAACCTGC 1692  
QY 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580  
DB 1693 ATGGAGTCTTTTCTGGAAGCGCAAGCTTGATGAGCCCAAGTATCTACCCGACACCTGGTG 1752  
QY 581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600  
DB 1753 TTACTGCAAGCGGCTGCTGATGGCTGGAGACGATCATGGTGCAGAAATTTGTGATCTA 1812  
QY 601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeu 620  
DB 1813 GGAGCAATTGACATGACTCTGGCAAGCGTGGCCACTTGGTGTGACAGCAGCTGGAAACTG 1872  
QY 621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640  
DB 1873 CAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTACTTGGAGACAAAGGCGCTTCTCTCAC 1932  
QY 641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnProPro 660  
DB 1933 GGCAACGCTCTCAGCACGGAAGGTGCTCTGGCTGCTGAGGGGGGTGATGGGAATCCACCT 1992  
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680

DB 1993 TTCATTAAAGCTGAGTGATCCTGGTCTCAGTCCCACTGTGCTGAGCCTGGAATGCTCACC 2052  
QY 681 AspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700  
DB 2053 GACAGAAATACCTGGTGGTGGCCCCCAATGTCTCCAGGAGGCTCAGACACTCTGCTTGGAG 2112  
QY 701 AlaAspLysTyrGlyPheGlyValAlaThrThrTyrGluValPheGlnArgGlyProAlaHis 720  
DB 2113 GCTGCAAGTGGGCTTTGGAGCCACCGTGGGAGGTGTTCCAAACGGGGAACCGCCCCAC 2172  
QY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlnGlnLeuPro 740  
DB 2173 ATCACCCTCGCTGGAGCCGCCCAAAAGCTGAAGTTCTATGAGACACGAGGACAGCTGCC 2232  
QY 741 AlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
DB 2233 GCTCTCAAATGACAGAACTGGCGGACTTATCACACAGTGCATGGCGTATGATCCTGGC 2292  
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780  
DB 2293 CGGCGCCCCCTCTTCGAGCTATCCTCAGAGACCTCAACGGCTCATTTACATCAGATTAC 2352  
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800  
DB 2353 GAGCTCCTCTCAGACCCACACCTGGCATCCCGAGTCCCGAGATGAGCTGTGCGTAGCT 2412  
QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyr 820  
DB 2413 GCGGCCAGCTCTATGCTGTCAGACCCCGCCATATTCGAGGAGAGACACCTTAAGTAC 2472  
QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840  
DB 2473 ATCTCTTTGCTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTGCCGTATGACCCCTG 2532  
QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860  
DB 2533 GGGACAATACGGGACCCCTGGTGGCAGTGAACAGCTACAGCACAGCGTGCCACACAG 2592  
QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880  
DB 2593 CAGAGGAGCTTCCAGCGGGAGATTGAGTCTTAAGGCTCTGCACAGCGACTTCATCGTC 2652  
QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900  
DB 2653 AAGTACCGGGGAGTCAGCTATGGGCGCAGTCCGACAGACCTGCGGTTTGGTATGAGTAC 2712  
QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGlyLeuHisThrAspArg 920  
DB 2713 TTGCCACGGCTGCTCGAGACTTGTGCGAGCGCATCGCGCTTGCACACCGACCGC 2772  
QY 921 LeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCys 940  
DB 2773 CTACTGCTTTGCTGTCAGATCTGCAAGGCGCATGGAGTACCTGGGTGCGCGCGCTGC 2832  
QY 941 ValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLysIle 960  
DB 2833 GTACACCGTGACCTGGCTGCGCGCAACATCTTGTGGAGAGCGAGGCTCATGTGAAGATC 2892  
QY 961 AlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGlu 980  
DB 2893 CGGACTTCGCGCTCGCTAAGCTGCTGCCCTCGGAAAGGACTACTAGCTGGTGGCGAG 2952  
QY 981 ProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAsnIlePheSer 1000  
DB 2953 CCGGCAAAAGCCCATCTTTTGTATGCCCCGGAGTCCCTATCTGACAAACATCTTCTCC 3012  
QY 1001 ArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAsp 1020  
DB 3013 CGCCAAATCTGAGCTGGAGTTTCGAGTGTGTGTGACGAGCTCTTCCACTACTGCGAC 3072  
QY 1021 LysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyPro 1040



QY	280	rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIle	299	Db	1951	TCAGGCAACGCTCTCAGCACGGAAGTGTCTCTGGCTCTGAGGGGGTGTATGGGAATCC	2050
Db	911	CGGGACACGAGGCTTTTCCAGACCTTCTGTGACTTTCCGGAAATCGTGGATGTCCAGCAT	970	QY	659	oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetIe	679
QY	299	eAsnGlnAlaProArgValGlyProAlaGlyCluHisArgLeuValThrValThrArgMe	319	Db	2051	ACCTTTTCATTAAAGCTGAGTGATCTCTGGTGTCTCAGTCCCACTGTGCTGAGCCTGGAAATGCT	2110
Db	971	CAAGCAGGCCCCACGCTGTGGGTCCGGCAGGGGAGCACCCTGGTGTCACTGTCCACAGGAT	1030	QY	679	uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysIe	699
QY	319	tAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAla	339	Db	2111	CACCGACAGAATACCTCTGGTGGCCCCGAATGTCTCCAGGAGGCTCAGACACTCTGCTT	2170
Db	1031	GGACGGCCACATCTGGGAAGCGAGTTTCGGGGGTCTGCTGAGGCGCTGTCTTCGTGGC	1090	QY	699	uGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl	719
QY	339	aLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVa	359	Db	2171	GGAGGCTGACAAAGTGGGGCTTTGGAGCCACACGCTGGGAGGTGTTCAGCGGGGACCCGC	2230
Db	1091	CCTCGTGATGGTACTTCCGCTCATCTGGACTCCAGGCATTTATTTCTGCAAGAGGT	1150	QY	719	aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLe	739
QY	359	lAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAs	379	Db	2231	CCACATCACCTCGCTGGAGCCGCCAAAAGCTGAAGTTCTATGAGGACCAAGGACAGCT	2290
Db	1151	GGCCGCCACCGCTGTGGAGGAGGAGCGGAGCTGTGCCATGGACCCATCACGTTAGA	1210	QY	739	uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr	759
QY	379	pPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuAtqAr	399	Db	2291	GCCCGCTCTCAAAATGGACAGAACTGGCGGAGCTTATCACAGATGCATGGCGTATGATCC	2350
Db	1211	CTTTGCCATCCACAGCTGAGGCGCTGGCTCCCTCCAGCACCTATATTTCTCCGCCG	1270	QY	759	oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs	779
QY	399	gSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPr	419	Db	2351	TGGCGGGCGCCCTCTCTCCAGAGTATCTCCAGAGACTCAACCGGCTCATTTACATCAGA	2410
Db	1271	CAGCCCGCAGGACTATGACAGCTTCTTCTTACCGCTGGCTGCAGACTCTCTTTGGCCC	1330	QY	779	pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVa	799
QY	419	oAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLe	439	Db	2411	TTACAGACTCTCTCAGACCCACACCTGGCATCCGAGTCTCCGAGATGAGCTGTGC--	2468
Db	1331	CGACTAACAGGGCTGCTCATCTCCGCCAGGACCCAGCGGGCTTTCTCCCTGGTGGGCT	1390	QY	799	lAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLy	819
QY	439	uSerGlnProHisArgSerLeuArgGluLeuAlaAlaCysTrpAsnSerGlyLeuAr	459	Db	2469	-GGTGGCGCCCGAGCTCTATGCTCTCCAGGACCCCGCCCATATTCAGGAGAGACACCTTAA	2527
Db	1391	CAGCCAGCCCCACAGAGCTTGGGGAGCTGCTGCAGCCTGCTGGAATCTTGGGCTGCG	1450	QY	819	sTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPr	839
QY	459	gValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSe	479	Db	2528	GTACATCTCTTGTGGCGAGGCGCACTTTGGCAGCGTGGAGCTGTGCCCTATGACCC	2587
Db	1451	AGTAGACGGTCTGCCCTGAACCTTAACATCTGCTGGCTCCAGACCCCAAGGAAAGTC	1510	QY	839	oLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAs	859
QY	479	rAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSe	499	Db	2588	CCTGGGGGACAAATACGGGACCCCTGTGGCAGTGAACAGACTACAGACACAGCGGGCCAGA	2647
Db	1511	CAATTTTATGCTGTGGTGGAGGGGCTGCACCCCGCGGCTCTGCCCTGTGCTCCCGCTC	1570	QY	859	pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIl	879
QY	499	rCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHi	519	Db	2648	CCACAGAGGAGACTTCCAGCGGGAGATTCCAGATCCTTAAAGCTCTGCACAGCGACTTCA	2707
Db	1571	CTGCTGTGGCTGACACAGCTGAGCTTCCACAAATCCACGAGACAGCTGGAGTGGCA	1630	QY	879	eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGl	899
QY	519	sGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVa	539	Db	2708	CGTCAAGTACCGGGAGTCCAGTATGGCCAGGTCGCCAGAGCCTGGTGGTGGATGGA	2767
Db	1631	CGAGAACTGGGTCACTGTTTATCAAGATCTTCCGTGGCGCAGCGGGAGGTCTGT	1690	QY	899	uTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTh	918
QY	539	lAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAs	559	Db	2768	GTACTCTCCACGGCTGCTGCGAGACTTCTCTGAGCGGCTCTGCGCGGCTGCACAC	2827
Db	1691	GGATGGTGAGACACATGACTCGGAAGTCTCTCTGAAGTCTATGAGCTCCAGACATCGGA	1750	QY	918	rAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetClnTyrLeuGlyAlaAr	938
QY	559	nCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLe	579	Db	2828	CGACCGCTACTGCTGTCTTGGCTGGCAGATCTGCAAGGGATGGAGTACCTGGTGGCGG	2887
Db	1751	CTGCATCGAGTCTTTCTGGAGCGGAGCTTGATGAGCCAAAGTATCTTACCCGACCT	1810	QY	938	gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVa	958
QY	579	uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTy	599	Db	2888	CCGCTGGGTACACCGTACCTGGCTGGCGCGCAACATCTTGGTGAGAGCGAGGCTCATGT	2947
Db	1811	GGTGTACTGCACGGCTCTGCATGGCTGGAGACAGCATCATGGTGCAGGAATTTGTGA	1870	QY	958	lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa	978
QY	599	rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLy	619	Db	2948	GAAGATCGCGGACTTCCGCCCTCGCTAAGCTGTGCCCTCGGAAAGGACTACTACGTTGT	3007
Db	1871	TCTAGGAGCAATTCACATGTATCTCCGCAAGCGTGGCCACCTGGTGTGTCAGCCAGCTG	1930	QY	978	lArgGluProGlyClnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIl	998
QY	619	sLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPr	639	Db	3008	CCCGAGCTTGGGCCAAAGCCCCATCTTTTGGTATGCCCGGAGTCCCTATCTGTACACAT	3067
Db	1931	ACTGAGGTGACCAAGACGCTGGCATATGCTTAACTTACTTTGAGGACAAAGGCCCTTC	1990	QY	998	ePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrClnLeuPheThrTy	1018
QY	639	oHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnPr	659				

```

Db      3068  CTTCTCCCGCAATCTGACGCTGCGAGCTTCGAGTGGTGTGTACGAGCTCTTACACCTA 3127
Qy      1018  rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgG1 1038
Db      3128  CTGCGACAGAGCTGACGCCATCCGCTAGTTCCTGCGCATGATGGGCGCTGAGCGTGA 3187
Qy      1038  uGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProPr 1058
Db      3188  AGACACCCCGCTCTCGCGCTCTCTGAGCTGTGGCAGAGGCCGACGCTCCACACCAC 3247
Qy      1058  oProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTTPAlaProGluProHi 1078
Db      3248  TCCACCTCCCGCCACCGAGTTCAGGAGCTCATGACGCTGTGCTGGCGCGCCAGCCGCA 3307
Qy      1078  sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTTPArgGlyArgPr 1098
Db      3308  CCACCGGCCAGGCTTCGGACCCCTGAGCCCGCCAGCTGGACGCGCTGTGGCGTGAAGACC 3367
Qy      1098  oGly 1099
Db      3368  CGGA 3371

RESULT 3
LOCUS   RATJAK3
DEFINITION Rattus norvegicus JAK3 3778 bp mRNA linear ROD 10-JAN-2003
ACCESSION D28508
VERSION D28508.1 GI:485811
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1
AUTHORS Takahashi, T. and Shirasawa, T.
TITLE Molecular cloning of rat JAK3, a novel member of the JAK family of
JOURNAL protein tyrosine kinases
MEDLINE FEBS Lett. 342 (2), 124-128 (1994)
PUBMED 94192816
REFERENCE 2 (bases 1 to 3778)
AUTHORS Shirasawa, T.
TITLE Direct Submission
JOURNAL Submitted (10-FEB-1994) Takuji Shirasawa, Tokyo Metropolitan
Institute of Gerontology, Department of Molecular Pathology; 35-2
Sakae-cho, Itabashi-ku, Tokyo 173, Japan
(TEL:81-3-3964-3241(ex.3034), FAX:81-3-3579-4776)
FEATURES
source
1..3778
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="JAK3"
/tissue_type="spleen"
/dev_stage="adult"
1..3778
/gene="JAK3"
61..3363
/gene="JAK3"
/codon_start=1
/evidence="experimental"
/product="protein-tyrosine kinase"
/protein_id="BA005868.1"
/db_xref="GI:485812"
FGDYLABDLCVRAKACGILPVHSLFALATEDLSWFPESHIFSIDVDTQVLVRL
/translation="MAPSEETELISQSCSLSSSEAGALHVLPRGPQPORLFS
RFPVPGWGLETCRFGHLKDLTSAILDVHLEHLPACHRSDLVSGRLPVGLSLKDOG
EFVLALDLAQARKAQRPGELLKSVKACLPSPRLDIOGQSVTRRRIRRTVV
QALAPCSLSRPAVMKAYILDLERLHPAATTESFLVGLPGQSEFGCLRVTGNGI
AWSKQDLFTQCDPFEIVDSIKQAPRVGPAGEHRLVTITRMDGHLEAPFGLPE
ALSFLVVDVGFRLICDSRHFFCKEVPAPPELLEBEEAELCHGPIITLDFAIHKLKAAGSL

```

```

PSSYLRRSPQDYSFLITACVQTPLEGDPYKCLIRQDPDSGAFSLVGLSOLHRSLOEL
LTACWHSGLQVDNALNTSCCPRPEKESNLIVVRGRNPTTPAPSHSSCCALTKLS
PHTIPADSLWHENLHGHSFTKIPHGREVVVDGTHDTEVLLKYMDSHQNCMESFL
EASLMSQSVSPHVLVLLHGVCMAGDSIMVOEYVLGAIDTYLRKRGHLVPASWKLQVT
KQLAYALNLEDKGLPHGNVSARKVLLAREGVGNPPRIKLSDPGVSPVLSLEMLTD
RIPWAPCELEAGLNLLEADKMGFGATWTFSGAPPHITSLLEPAKKLKFYEDRGQL
FALKWTEGLEGLTACMAYDPPRRPSFRALIRDLNGLITSDYELLSDPTPGIPNPROEL
CGGAQYACQDPATFEERHLKYISLLKGNFGVSLCRYDPLGMDNTGPIVAVKQJHS
GPEQORDFREIQILKHCDFIVKRGVSGPGRQELRLVMEYLPSGCLRDFLQHRH
ARLHNDRLLPWAOICKMEYLKAGRCVHRDLAARNILVSEHVKIADFGAKLLPL
GKDYVVRPGSPPIFWAPESLSDNIESRSDVMSFGVLYELFTYSTSKSPSTEF
LKMIGPERGSPGLCHLLELLEAGRRLLPPPTCTCTEVEQLMLQCNENPQDRPAFTLS
PQLDALWRGSPG"
BASE COUNT 729 a 1172 c 1060 g 817 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3778
Score: 5380.50 Matches: 1018
Percent Similarity: 94.64% Conservative: 24
Best Local Similarity: 92.46% Mismatches: 56
Query Match: 91.82% Indels: 3
DB: 10 Gaps: 3
US-09-397-967A-16 (1-1099) x RATJAK3 (1-3778)
Qy 1 MetAlaProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20
Db 61 ATGGCACCTCCAAAGCGAGGAGCGCTCTCACTCTCAGGGCTCTTAGCCTCTCATCC 120
Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 121 TCAGAAAGCAGGAGCCCTGCATGTCTCTCTCCCGGGGACCTGGCGCTCCCGACGGA 180
Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 181 CTGTCACTCTCTTTGGGACTACTTGGCTGAGGATCTGTGTGTGGGGCTGCCAAGGCC 240
Qy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 241 TGTGGCATCTGCCCGTATTATCATTCGCTCTTCCCTCGCCACCGAGGACTTGTCTTGC 300
Qy 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 301 TGGTTCCTCCCAAGCACACATCTTCTCATAGAGGACGTGGACACTCAAGTCTTGGTCTAC 360
Qy 101 ArgLeuArgPheTyrPheProAspTTPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 361 AGCTACGCTTTATTCTCTGTTGTTGGACTGGAGACCTGTACCCGCTTTGGGTGA 420
Qy 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 421 CACAAGATTGGACAGTGCCATCTTTCAGTACATGTTTTAGAACACCTCTTTGCCAG 480
Qy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 481 CACCGCAGTGACCTGTGTAGTGGCGCTCTCCAGTAGGCGCTTAGCCTGAAGGACCCAGGA 540
Qy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 541 GAGTTCCTGAGCCTAGCTGTGTGGACCTGGCCAGATGGCTGTAGAGCAGGCCACCGCG 600
Qy 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 601 CCAGGAGAGCTGCTGAAGTCTGTCTAGTACAAAGCCTGTCTGGCGGCCACGCTGCGCGAC 660
Qy 201 ValIleGlnGlnAsnPheValThrArgArgArgIleArgThrValValLeuAla 220
Db 661 CTGATCCAGGCGCCAGAGCTCTGTGACGCGGAGGCGCATCCGACGAGCCGTGGTCCAGCGC 720
Qy 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240
Db 721 CTGGCGCGGTGTAGTACCTTGCACGCGGACCCCTACGCACTCATCGGCCAGTACATCTTA 780

```



241 AspLeuGluArgLeuHisProAlaThrThrGluThrPheArgValGlyLeuProGly 260  
 781 GACCTTGAGCGCTGCATCCAGCGCCACCATGAGTCTTCTCTGTGGAGATCCCGGGC 840  
 261 AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyLeuProTrpSerSer 280  
 841 GCCAGGAGGAGCGGGTTGCGTGTGAGCAGGGGATAATGGCATCGCTGGAGCTCC 900  
 281 AsnAsp---GluLeuPheGlnThrPheCysAspPheProGluLeuValAspValSerIle 299  
 901 AAGGACCGAGGAGCTTTCCAGACCTTCTGTGACTTTCCGAAATCGTGATGTACGATC 960  
 300 AsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMet 319  
 961 AAGCAGGCTCCACGCGTGGTCCGAGAGGGAGCACCAGCTAGTCACCATCACAGGATG 1020  
 320 AspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAla 339  
 1021 GACGGTCACATCTCTGAAAGCGAGTTCCGGGGCTGCTGAGGGCTGTCTTTCGTGGCG 1080  
 340 LeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVal 359  
 1081 CTCGTGGATGGGACTTTTCGCGCTGATCTCGCACTCCAGGCATTTCTTCTGCAAGAGGTG 1140  
 360 AlaProProArgLeuGluGluAlaAspValCysHisGlyProIleThrLeuAsp 379  
 1141 GCACCGCCACGGCTGCTGGAGAGGAGCGAGCTGTGCCATGGACCCATCACGTTAGAC 1200  
 380 PheAlaIleHisLeuLeuAlaGlySerLeuProGlyThrTyrIleLeuArgArg 399  
 1201 TTTGCCATCCACAAGCTGAAGGCTGCTGGCTCCCTCCCGGGCTCTACATTCTCCGACGC 1260  
 400 SerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPro 419  
 1261 AGCCCGCAGGACTATGACAGCTTCTTCTTACTGCTGTGTCCAGACTCTCTTGGCCCC 1320  
 420 AspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeu 439  
 1321 GACTACAGGGCTGCCTCATCGCCACAGACCCAGTGGGGCTTCTCCCTGTGGTCTC 1380  
 440 SerGlnProHisArgSerLeuArgGluLeuAlaCysTrpAsnSerGlyLeuArg 459  
 1381 AGCCAGCTGCACAGAAGCTTACAGGAGCTGTCTACAGCCTGCTGGCAITCTTGAGTGC 1440  
 460 ValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSer 479  
 1441 GTAGACGGCACTGCCCTGAACCTCACATCTTGTCTGCTCCCGCAGACCCCAAGGAAAGTCC 1500  
 480 AsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSer 499  
 1501 AATTGTGTCGTGGTGTAGGGCGCGCAACCCAGCTGCCCGCGGCACTCCCGCATCC 1560  
 500 CysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHis 519  
 1561 TGCTGTGGCTGACTAAGTTGAGCTTCCACAAATCCCGCAGACAGCCTGGAGTGGCAC 1620  
 520 GluAsnLeuGlyHisGlySerPheThrIlyIlePheArgGlySerArgArgGluValVal 539  
 1621 GAGAACTTGTGTCAAGCTCTTTTACCAGAGATCTTCCATGGCCATAGCGGGAGGTTGTG 1680  
 540 AspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsn 559  
 1681 GATGGTGACACATGACAGAGTCTCTCTGAAGGTTATGGACTCCAGACACAGAAAC 1740  
 560 CysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeu 579  
 1741 TGCATGGAGTCTTCTTGGAGCGCAGCTTGTATGATGCCCAAGTATCTTACCCGACCTG 1800  
 580 ValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyr 599  
 1801 GTGTTGTGTCATGGAGTGTGATGGTGGAGACGATCATGGTGCAGGAATTTGTGTAC 1860

600 LeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLys 619  
 1861 CTTGGAGCAATTTGACACGTACTGCGAAGCGTGTGTACCTGGTCCGCCAGCTGGAAA 1920  
 620 LeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPro 639  
 1921 CTGCAAGTGACCAAGCAGCTGGCTATGCCCTTAATCTACTTGGAGACAAAGGCCCTCCCT 1980  
 640 HisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnPro 659  
 1981 CATGCAACAGCTCTACGCGAGGAAGGTGTCTAGCTCGTGGGGGTTCATGGGAATCCA 2040  
 660 ProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeu 679  
 2041 CTTTTCATCAAGCTGAGTGATCTCTGTGTACGCCCAACTGTTTTCAGGCTGGAAATGCTC 2100  
 680 ThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699  
 2101 ACTGACAGGATTCCTGGTGGCCCCGGAGTGTCTCCAGGAAGCTGGAAACGCTCAACTTG 2160  
 700 GluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAla 719  
 2161 GAGGCTGCAAGTGGGGCTTTGGCGCCACACGCTGGAGGTGTTTCAGCGGGCACCCCATG 2220  
 720 HistIleThrSerLeuGluProAlaLysLysLysLysPheTyrGluAspGlnGlyGlnLeu 739  
 2221 CACATTACTCGTGGAGCCCGCCAAAGAGTGAAGTTCTATGAGGACCGGGGACAGCTG 2280  
 740 ProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759  
 2281 CCGGCCCTCAATGACAGAGCTGGAGGGGTCTATCGCACAGTGCATGGCGTATGATCTCT 2340  
 760 GlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779  
 2341 GCGCGGCGCCCTCTCTCCGAGCCATCTCCAGAGACCTCAACGGCCTCATTTACATCAGAT 2400  
 780 TyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVal 799  
 2401 TACGAGTCTCTCTCAGACCCACACCTGGCATCCGAAATCCCGAGATGAGCTGTGT--- 2457  
 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys 819  
 2458 GGTGGTGCCAGCTCTATGCTCGCAGACCCCGCCCATCTTCGAGGAGAGACACCTTAAG 2517  
 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro 839  
 2518 TACATCTCCCTGCTGGGCAAGGGAACTTCGGCAGCGTGGAGCTGTGCGCTACGACCT 2577  
 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859  
 2578 CTGGAGACAATACTGGGACCCCTCTGTAGCAGTGAAGCAGTTTACAGCACAGCGGGCCAGAA 2637  
 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879  
 2638 CAGCAGAGGACTTCAGCGGAGAGATTCAGATCTCTCAAGGCTCTACACTGTGACTTCACT 2697  
 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899  
 2698 GTCAAGTACCGGGAGTCAGCTATGGGCTGTCTCCGAGGAGCTGCGGTAGTGTAGTGAG 2757  
 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisThr 918  
 2758 TACTGCGCCAGCGCTCGCTCGGGACTTCTCTCAGCGCCACCGCGCGCGCTCTCACAAC 2817  
 919 AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyValArg 938  
 2818 GACCGCCTACTGTGTGTGCTGGCAGATCTGCAAGGGCATTTGGAGTACCTTGGGGCGACGC 2877  
 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal 958  
 2878 CGCTGCGTCATCGTGAACCTGGCTGCGCGCAACATCTCTGGTGGAGAGCTGAGGCGCATGTG 2937  
 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978

```

Db      2938  AAGATTGCTGACTTGGCGCTCGCTAAGCTGCTGCCTCTTGGAAAGACTACTACGTGGTC 2997
QY      979   ArgGluProGlyGlnSerProIlePheThrTyAlaProGluSerLeuSerAspIle 998
Db      2998  CGCGTGGCGGGCCAAAGTCCCATCTTCTGATAGCGCCCTGAGTCCCTATCTGACACATC 3057
QY      999   PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db      3058  TTCTCCCGCAATCTACGTGTGAGCTTGGCGGTGGTCTGTATGAGCTCTTCACCTAC 3117
QY      1019  CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038
Db      3118  ACGGACAAAGCTGCAGCCCATCTGAGTTCCTCGGTATGATTGGCGCTGAGCGTGAG 3177
QY      1039  GlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProPro 1058
Db      3178  GGATCCCGCTCTGCACCTCTTGGAGCTGCTGGCAGAGCGCGCTCTCCACACCC 3237
QY      1059  ProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHis 1078
Db      3238  TCCACTGCGCTTACTGAGGTTCAAGAGCTCATGCAGCTGTGCTGGCCCAACCCACAA 3297
QY      1079  AspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPro 1098
Db      3298  GACCGGCGCGGCTGTACACCTTGAGCCCGCCAGCTGATGCGCTGTGGCGCGGAGCCCC 3357
QY      1099  Gly 1099
Db      3358  GGA 3360

```

```

RESULT 4
MUSJAK3H
LOCUS      MUSJAK3H               4016 bp      mRNA      linear      ROD 13-SEP-1994
DEFINITION Mus musculus (clone 32DS-1) protein tyrosine kinase (JAK3) mRNA.
ACCESSION  L33768
VERSION    L33768.1  GI:508533
KEYWORDS   protein tyrosine kinase.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Mammalia; Eutheria; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS    Rane, S.G. and Reddy, E.P.
TITLE      JAK3: a novel JAK kinase associated with terminal differentiation
            of hematopoietic cells
JOURNAL    Oncogene 9 (8), 2415-2423 (1994)
MEDLINE    94309920
PUBMED     7518579
COMMENT    Original source text: Mus musculus cDNA to mRNA.
FEATURES   Location/Qualifiers
            1..4016
             /organism="Mus musculus"
             /mol_type="mRNA"
             /db_xref="taxon:10090"
             /clone="32DS-1"
             /cell_line="32Dcl3"
             /tissue_lib="lambda gt11"
            7..3906
             /note="protein tyrosine kinase"
             /codon_start=1
             /product="JAK3"
             /protein_id="AAA21415.1"
             /db_xref="GI:537607"
            /translation="MAPPSSETPLIPQSCSLSSSEAGALHVLPPRGPPQPLSPS
            FGDLAEDLCVRAKACAIILPVHVSFALATEDFSCWFFPSHIFCIEDVDTVLVYRL
            RFYPPDFGLETHRFGRKDLTSAILDLHLVLFQAHSRLVSGRLPVGLSKMEQG
            EFLSLVDLIAQMARQOAPGELLKTVSKACILPPLSLRDVIQQNPVTRIRRTTV
            LALRVSPARTATAGQVYLDLRLKHPAATTETFRVGLPQAOEPLRLVAGDNGLS
            WSSDQGLGLSLRGVRGAWRLVKPVTIRKVASPGRADCTTGCGGGLNKLKVPQ
            MELPQGLTGWTRVLRDGRGTEGSDMWISGHDTRTPVPSLTSPPPHKWRWG
            GRGCGAGSRVLPWLSLFLFFFNFGARQFSYSSGCGPTFRVPRGWPTQKSAQL
            CLSSAVIKGRVPLRLYCLSLPQLPQFCDFPEIVDVSIKQPTCGSGRHLVTVTRM

```

```

misc_feature 2199..3024
            /note="JAK homology domain II (JH2)"
misc_feature 3076..3876
            /note="JAK homology domain I (JH1)"
BASE COUNT 742 a 1203 c 1235 g 836 t
ORIGIN
Alignment Scores:
Pred. No.: 3 05e-311      Length: 4016
Score: 5223.00           Matches: 1060
Percent Similarity: 80.59%      Conservative: 3
Best Local Similarity: 80.36%    Mismatches: 32
Query Match: 89.13%           Indels: 235
DB: 10                     Gaps: 6
US-09-397-967A-16 (1-1099) x MUSJAK3H (1-4016)
QY      1   MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db      7   ATGGCACTCCCAAGTCAGGAGACACCTCTGATCCCTCAGCGCTCTTGCGACCTCTCATCC 66
QY      21   SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db      67   TCAGAGCGAGGAGCCCTGTCATGTGCTCTCTCCCGGGGACCTGGCGCTCCCGACGGA 126
QY      41   LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db      127  TTGTGATCTCTTTTGGGACTACTTGGCTGAGGATTTATGTGCGAGCTGCCAGGCC 186
QY      61   CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db      187  TGTGCGATCTGCTGCTGTTTATCATCTGCTTTTCCCTCTGCGCCTGAGGACTCTCTTTC 246
QY      81   TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db      247  TGGTTCCTCCCAAGCCACATCTTCTGATAGAGAGCGTGGACACTCAAGTCTTGGTCTAC 306
QY      101  ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db      307  AGGCTACGCTTTTATTTCTCTGCTGCTGTTGGCTGGAGCATGTACCTGCTTGGGCTG 366
QY      121  ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db      367  CGCAAGATTGACCACTGCTGCTTACATGCTTTAGAACATCTCTTGTCTCAG 426
QY      141  HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db      427  CACCGCAGTACCTGTTGAGTGGCGCTCCCGGTGGCGCTTAGCATGAAGGAGCAGGGA 486
QY      161  GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db      487  GAGTTCCTGAGCTGGCGCTGCTGGACTTGGCCAGATGGCTCTGAGCAGGCCAGCGC 546
QY      181  ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db      547  CCAGGAGAGCTGCTGAAGACGGTCAGTTACAAAGCCTGTCTGCGCGCCAGCCTCGCGAT 606
QY      201  ValIleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAla 220

```

```

DCHLEAFPPGLPEALSFVALVDGYFRILCDRHRHYCKEVAAPRLLEEAEELCHGPIT
LCASQGMGMAAARTVLGWTWLSWPLDFAIHKLKAAASLPGTYILRRSPQDYDS
FLITACVDTPLGDKYKGLIRQDPSGASFLVASAPTEACSCLOLEFGLRVDGAAL
NLTSACAPRPEKSNLIVVRGCTPAPAPGSPSCALTLQSFHTIPTDLSLGHENLGH
GSFTKIFGRRRRNVVDGTHSEVLLKYMDSHRNCMESFLKAAASLVMSOVSYPHLVLL
HGVCMAGDSIMVQEFVYLGADIMYLKRKGLHVSWKLOVTKQLAYALNYLEDKGLPH
GNVARKVLLAREGGDGNPFIKLSDPGVSPITVLSLEMLTDRIFWVAPCELOEATLCL
LEADKMGATATWTFVSGPAHITSLEPAKLKYFEDDQQLPALKWTEAGLITQCMAL
YDGRPPFRATLRDLNGLITSYDLNTGPHLASRVLEMSCAWRPALCLPGRHIEER
HLKYSLLKGNFGVSELRCYDPLNTGFLVAVKLOHSGPQQRDFQREIQILKALH
SDFIVKRGVSGPGRQSLRLVMEVLPSCGLDFLQRHAAALHTDLRLLFAMQICKGM
EYLGAARRCHRDLAARNLIVSEAHVKIADFGLAKLPLCKDYVYVREPGSQPIFWYA
PELSDNITFSRQSDVMVSGVLYELFTYCDKSCSPSAEFLRMWGPEREPEPFLCRLLEL
LAEGRRLLPPPTCTPTTEVQELMQLCVAPSRRTTQPSAP"
misc_feature 2199..3024
            /note="JAK homology domain II (JH2)"
misc_feature 3076..3876
            /note="JAK homology domain I (JH1)"
BASE COUNT 742 a 1203 c 1235 g 836 t
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3 05e-311      Length: 4016
Score: 5223.00           Matches: 1060
Percent Similarity: 80.59%      Conservative: 3
Best Local Similarity: 80.36%    Mismatches: 32
Query Match: 89.13%           Indels: 235
DB: 10                     Gaps: 6

```

US-09-397-967A-16 (1-1099) x MUSJAK3H (1-4016)

607 GTGATCCAGGCCAGAACTTCGTGACACGACGCGCGATCCGACGACCGTGGTCTTGGCG 666  
Db  
221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240  
Qy  
667 CTGC--GCGTGTGTGCTCCAGGCCGCGCTACGCGTCATGCGCAAGTATAT-CTG 723  
Db  
241 AspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly 260  
Qy  
724 GACCTGGAGCGCTACATCCAGCGGCCACACCGAGACCTTCGCTGGGGCTCCCGGGC 783  
Db  
261 AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer 280  
Qy  
784 GCCCAGGAGGAGCGGGGCTTCCTGGTGTGGCGGGGACAAACGGCATCTCCCTGAGCTCC 843  
Db  
281 AsnAsp----- 282  
Qy  
844 GGGACACGAGGTGTGGGCTTGGGCTTGAAGAGCGGGGTCCGCGGGGAGCGGTGGCGG 903  
Db  
282 ----- 282  
Qy  
904 TTAGTAAGAAGCCCGTCCGAACAATCAGGAAGTGGCTAGTCCCGGGCGGCTGATTGC 963  
Db  
282 ----- 282  
Qy  
964 ACTACGGGCGAGGGCGGGGTCAAATTTAAAGGTGGGTCCGGGATGGAGCTGCCCCAG 1023  
Db  
282 ----- 282  
Qy  
1024 GGACTGACCTGGGCGGTGACTAGAAGGTGAGACTAGATAGGGGGCGGGAGCGACAGAA 1083  
Db  
282 ----- 282  
Qy  
1084 GGGATAGTATGATTGATCAGTGGGCATGACCCGACCGAGCCCTGTCTTCAGTCCCCCTA 1143  
Db  
282 ----- 282  
Qy  
1144 ACCTCGTCCCCCCCCCAAGTGGAGTGGGAAGGAGGAGGAGGTGTGCT 1203  
Db  
282 ----- 282  
Qy  
1204 GGGTCAAGGTCTGAATCCCTGGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1263  
Db  
282 ----- 282  
Qy  
1264 TTTCCGAGACAGGGTTTCTCGTATAGCTCTGGCTCTCTGGAACTCATTGTAAGACCA 1323  
Db  
282 ----- 282  
Qy  
1324 GGCTGGCCTCGAACTCAGAAATCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1383  
Db  
283 ----- 282  
Qy  
1384 CGTGTGCCACTACGCGCTACTGCTGTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1443  
Db  
290 AspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGly 309  
Qy  
1444 GACTTTCGGAAATCGTGGATGTCAGCATCAAGCAGC--CACCTGTGGTCCGCGCAGG 1501  
Db  
310 GluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhePro 329  
Qy  
1502 AG-CACCGGCTGGTCACTGTCCAGGATGACGCCACATCTCTGGAAGCGGAGTTTCG 1560  
Db  
330 GlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys 349  
Qy  
1561 GGGCTGCTGAGGGCGGTCTCTTCGTGGCCCTCGTGGATGGGTACTTCCGCTGATCTGC 1620  
Db  
350 AspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluAla 369  
Qy  
1621 GACTCCAGGCATATTCTGCAAGAGAGGTGGCGGCCACCGCTGCTGGAGAGAGAGCG 1680  
Db  
370 AspValCysHisGlyProIleThr----- 377  
Qy  
1681 GAGCTGTGCCATGACCCATCACGTTAGGAGCTAGTGGGGTCAACAGTGGGGATGGGA 1740  
Db

378 -----LeuAspPhe 380  
Qy  
1741 TGGCAGCGGAGGACTGTCTTGGGTGGACCTGGCTGTGTCTTGGCCCTAGGTTAGACTTT 1800  
Db  
381 AlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgSer 400  
Qy  
1801 GCCATCCACAAGCTGAAGCGCGCTCCCTCCAGGACCTATATTTCTCCGCGCAGC 1860  
Db  
401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420  
Qy  
1861 CCGCAGGACTATGACAGCTTTCTTTACCGCTCCGCTCCAGACTCCTCTTGGCCCGAC 1920  
Db  
421 TyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440  
Qy  
1921 TACAAGGCTCCCTCATCCGACAGACCCAGCGGGGCTTTCTCCCTGGTTGC-CTCAGC 1979  
Db  
441 GlnProHisArgSerLeuArgGluLeuAlaAlaCysTrpAsnSerGlyLeuArgVal 460  
Qy  
1980 CAGCCCCACAGAAGCTGCGGAGCTGCTTGACG-TGCTGGAATTC-GGCTGGAGTA 2037  
Db  
461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsn 480  
Qy  
2038 GACGCTGTGCTGCTGAACCTAACATCTCTGCTGCGCTCCAGACCAAGAAAAGTCCAAT 2097  
Db  
481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500  
Qy  
2098 TTGATCGTGTGGAGAGGGCTGACCCCGCGCTGCCCCCTGCTGCTCCCGTCTTCG 2157  
Db  
501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu 520  
Qy  
2158 TGTGCGCTGACACAGCTGAGCTTCCACACAATTTCAACGGACAGCTG---GGACACGAG 2214  
Db  
521 AsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp 540  
Qy  
2215 AACCTGGTCACTGGTCTTTTACCAAGATCTTCGCTGCGCGCAGCGCGGAGTGTGGAT 2274  
Db  
541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560  
Qy  
2275 GGTGAGACACATGACTCGGAAGTCTCTGAAAGTCAATGAGCTCCAGACATCGGAACATGC 2334  
Db  
561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580  
Qy  
2335 ATGAGTCTTTTCTGGAAGCGCGCAAGCTTGATGAGCCAGTATCTTACCAGCCTGGTG 2394  
Db  
581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600  
Qy  
2395 TTACTGCACGCGCTCTGCATGGCTGGAGACAGCATCATGTTGTCAGGAATTTGTGTATCTA 2454  
Db  
601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisIleValSerAlaSerTrpLysLeu 620  
Qy  
2455 GGACAAATTTGACATGTACCTCGCGCAAGCTGCGCCCTCTGTTGTGAGCACTGGAACATG 2514  
Db  
621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640  
Qy  
2515 CAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTACTTGGAGGACAAAGGCTTCTCTAC 2574  
Db  
641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyAspGlyAsnProPro 660  
Qy  
2575 GGCAACGCTCTCAGCACGAAAGGTCTCTGCTGCTGAGGGGGGTGATGGAAATCCACCT 2634  
Db  
661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680  
Qy  
2635 TTCAATTAAGCTGAGTGTATCTGTGTGTCAGTCCCACTGTGTGAGCTGGAAATGCTCACC 2694  
Db  
681 AspArgIleProTrpValAlaProGluCysLeuGlnAlaGlnThrLeuCysLeuGlu 700  
Qy  
2695 GACAGAATACCTGGGTGGCCCCCGAATGCTCCAGAGGCTCAGACACTCTGCTTGGAG 2754  
Db  
701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720  
Qy  
2755 GCTGACAAATGGGGCTTTGGAGCCACACCTGGGAGGTGTTTCAGCGGGGAGCCCGCCAC 2814  
Db

QY 721 IleThrSerLeuGluProAlaIylsLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740  
Db 2815 ATCACTCTCGTGGAGCCGCCAAAGAGCTGAGTTCTATGAGGACACGAGCAGCTGCC 2874  
QY 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
Db 2875 GCTCTCAATGGACAGACTGGCGGACATTATCACACAGTGATGCGTATGATCTGGC 2934  
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780  
Db 2935 CGCGCGCCCTCTTCGAGCTATCTCAGAGACTCAACGGCCCTATTATCATCAGATTAC 2994  
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800  
Db 2995 GAGCTCTCTCA-GACCCACACTTGGATCCCGAGTCTTCGAGATGAGCTGTGGT---- 3049  
QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGlnArgHisLeuLysTyr 820  
Db 3050 GCGCGCCAGCTCTATGCTGCCAGGACCCCGCATATC-GAGGAGAGACACCTTAAGTAC 3108  
QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840  
Db 3109 ATCTCTTTTGTGGCAAGGCAACTTTGGCAGCGTGGAGCTGTGCCGTATGACCCCTG 3168  
QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860  
Db 3169 ---GACAAATACGGGACCCCTGGTGGCAGTGAACAGCTACAGCAGCGGCCAGACCAG 3225  
QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880  
Db 3226 CAGAGGAGCTCTCAGCGGAGATTAGATCTTAAAGCTCTGCACAGCGACTTCATCGTC 3285  
QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgIleValMetGluTyr 900  
Db 3286 AAGTACCGGGAGCTAGCTATGGCCAGGTCGCCAGAGCCTGCCGTGTGTGGAGTAC 3345  
QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg--GlyLeuHisThrAsp 919  
Db 3346 CTGCCACGCGCTCTCTCGAGACTTCTGCAGCGGCATCGCGGCCCTTGACACCGAC 3405  
QY 920 ArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArg 939  
Db 3406 CGCCTACTGCTGCTGCTGGCAGATCTGCAAGGCGATGGAGTACCTGGGTGGCGCGCC 3465  
QY 940 CysValHisArgAspLeuAlaArgAsnIleLeuValGluSerGluAlaHisValLys 959  
Db 3466 TGGGTACACGCTGACCTGGCTGGCGCAACATCTTGGTGAGAGCGAGGCTCATGTGAAG 3525  
QY 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArg 979  
Db 3526 ATCGCGGACTTCGGCCTCGCTAGCTGTGCTGCCCTGGGAAAGACTACTACGTGTGGC 3585  
QY 980 GluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIlePhe 999  
Db 3586 GAGCCTGGCCAAAGCCCATCTTTGGTATGCCCGGAGTCCCTATCTGCAACATCTTC 3645  
QY 1000 SerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCys 1019  
Db 3646 TCCCGCAATCTCAGCGTGGAGCTTCGGAGTGGTGTGTGACGAGCTTCTTCACTCTG 3705  
QY 1020 AspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGly 1039  
Db 3706 GACAAGAGCTGCAGCCATCCGTGAGTTCCTCGCATGATGGGCTGAGCGTGAAGA 3765  
QY 1040 ProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgLeuProProPro 1059  
Db 3766 CCCCGCTCTGCGCCCTCTCGAGCTCTGGAGCGGCGAGCGCTCCACCACCTCCC 3825  
QY 1060 ThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAsp 1079  
Db 3826 ACCTGCCCCACGAGGTTGAGAGCTCATGCACTGTGGTGGCGGCCAG-CGCGACGAC 3884  
QY 1080 ArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPro 1098

Db 3885 CGGCCAGCCTTCGGCACCTGAGCCCGACGCTGAGCGCTGTGGCGTGAAGACCC 3941  
RESULT 5  
AX203081 4064 bp DNA linear PAT 30-AUG-2001  
LOCUS Sequence 2 from Patent WO0152892.  
DEFINITION AX203081  
ACCESSION AX203081  
VERSION AX203081.1 GI:15392438  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Vassios,G.  
JAK/stat pathway inhibitors and the uses thereof  
TITLE Patent: WO 0152892-A 2 26-JUL-2001;  
JOURNAL GENZYME CORPORATION (US)  
FEATURES  
source  
1..4064  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 746 a 1292 c 1168 g 858 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,82e-278 Length: 4064  
Score: 4684.00 Matches: 905  
Percent Similarity: 87.61% Conservative: 64  
Best Local Similarity: 81.83% Mismatches: 129  
Query Match: 79.93% Indels: 9  
DB: 6 Gaps: 5  
US-09-397-967A-16 (1-1099) x AX203081 (1-4064)  
QY 1 MetAlaProProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
Db 96 ATGGCACCTTCAAGTAGAGAGAGCGCCCTGATCCCTCAGGTTTCATGACGCTCTGTGCC 155  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
Db 156 ACGGAGGCTGTGGCTGTCATGTGCTGCCCGCTCGGCGCCCGGGCGCCCGCCAGCGC 215  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
Db 216 CTATCTTCTCTCTTTCGGGACCATCTTGGCTGAGGACCTGTGCTGAGGCTGCAAGGCC 275  
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
Db 276 AGCGGAGATCTGCTGCTGTACCACTCCCTTTGTCTTGCCACGAGGACCTGTCTGTC 335  
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
Db 336 TGGTTCCTCCCGAGGACACATCTTCTCCGTGGAGATGCCAGCACCCCAAGTCTGTGTATC 395  
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
Db 396 AGGATTTCGCTTTTACTTTCCTCCCAATTTGGTGTGGGCTCGAGAAGTGCACCGCTTCGGGCTA 455  
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
Db 456 CGCAAGGATTTGCCAGTGTCTATCTTGTACCTGCCAGTCTCTGGAGCACCTCTTTGCCAG 515  
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
Db 516 CACCGCAGTGACCTGGTGTGGCGCCTCCCGTGGCGCTCAGTCTCAAGGAGCAGGGT 575  
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
Db 576 GAGTGTCTGAGCCTTGCCCGGTGTGGACCTTGCCCGGATGGCGGAGAGCGAGCCCGCGG 635



```

Db      2792  GCTGTGTCATGGAGTACCTGCCACGCGGCTGCTTGGCGAGCTTCTCGACGGCACCGCGC 2851
QY      915  y---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGlnTy 934
Db      2852  GCGCTCGATGCGACGCGCTCTCTTCTATTCTCGAGATCTGCAAGGGCATGGAGTA 2911
QY      934  rLeuGlyAlaArgAtgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
Db      2912  CTGGGCTCCCGCGCTGGTGACCGACCGACCTGGCGCGCCGAAACATCTCGTGAGAG 2971
QY      954  rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974
Db      2972  CGAGGCACACGTCACATCGCTGACTTCGCTGAGCTAGCTAAGCTGCTCGCGCTTGACAA 3031
QY      974  pTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLe 994
Db      3032  CTACTACGTGGTCCGGAGCAGCCAGCCAGCCCATTTCTTGATGTCGCCCGCAATCCCT 3091
QY      994  uSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGl 1014
Db      3092  CTGCGACACATCTTCTTCGCCAGTCAGACGCTTCGGGCTCGCTGTCGACGA 3151
QY      1014  uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGl 1034
Db      3152  GCTCTTACACTAGCGACAAAGCTGACGCCCTCGCGCGAGTCTCGCGGATGATGGG 3211
QY      1034  yProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgAr 1054
Db      3212  ATGTGAGCGGGATGTCGCCGCCCTCTGCGCCCTCTTGGAACTGCTGGAGAGGCGCAGAG 3271
QY      1054  gLeuLeuProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074
Db      3272  GCTGCGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3331
QY      1074  aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094
Db      3332  CCCTAGCCACAGACGCGCGCATCATTCAGCGCCCTGGCGCCCTGGCGCATGCTGTG 3391
QY      1094  pArgGlyArgProGly 1099
Db      3392  GAGCGGAAGCCGGGG 3407

RESULT 6
HSU09607
LOCUS      4064 bp      mRNA      linear      PRI 09-JUL-1994
DEFINITION Human JAK family protein tyrosine kinase (JAK3) mRNA, complete cds.
ACCESSION U09607
VERSION    U09607.1 GI:508730
KEYWORDS   .
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            (bases 1 to 1124)
AUTHORS    Kawamura,M., McVicar,D.W., Johnston,J.A., Blake,T.B., Chen,Y.,
            Lal,B.K., Lloyd,A.R., Kelvin,D.J., Staples,J.B., Ortaldo,J.R. and
            O'Shea,J.
            Molecular cloning of L-JAK, a Janus family protein-tyrosine kinase
            expressed in natural killer cells and activated leukocytes
            Proc. Natl. Acad. Sci. U.S.A. 91 (14), 6374-6378 (1994)
JOURNAL    MEDLINE 94294384
PUBMED     8022790
REFERENCE  2 (bases 1 to 4064)
AUTHORS    O'Shea,J.
TITLE      Direct Submission
JOURNAL    Submitted (11-MAY-1994) John O'Shea, Leukocyte Cell Biology Section
            LEI BRMP, National Cancer Institute FCRDC, Bldg 560 Rm 3146 FCRDC,
            Frederick, MD 21702, USA
FEATURES   location/Qualifiers
            1..4064
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"

```

```

/cell_line="YT, HUT-78"
/cell_type="natural killer cell, PHA activated T cells"
1..4064
/gene="JAK3"
96..3470
/gene="JAK3"
/codon_start=1
/product="JAK family protein tyrosine kinase"
/protein_id="AAA19626.1"
/db_xref="GI:508731"
/translation="MAPSEETPLIPORSCSLSTEAGALHVLIPARPGPQRLSFS
FDHLEADVOAKASGLIPVHSLFALATEDLSWFPFPHSHFVEDASTQVLYRI
RYFFNWFLEKCHRFGRKDLASAILDLVLEHLFAQHSDLVSRPLVSLKEQG
ECLSLADLARMAREQAPGELLKTSYKACLPSLEDLIQGLSFTRRAIRRTVR
RALPRVAACQADRHSIMAKYIMDLERLDPAGAAETPHVGLPGALGHDGLIRVAGD
GGIAWTCGQEVLPFCDEIVDISIKOAPRVGAGEHRLVTVTNDQILAEAPPK
LEASLFFVALVGVYFRLTDSOHFFCKEYVAPRLLIEVAEOCHGPTLDFANKLKTG
GRPGSVLRRSFQDFPDLFVVCVONPLGPDYKGLIRRSPTGFLVGLSRPHSSL
RELLATCMGGJLHVGVAVLTSCCIPRPEKSNLIVQGHSPPTSSILVQSQYQL
SMTFHKIPADSLWEHNIHGHSFTKIYRCRHEVVDGEARKTEVLKVMYDKKNCM
BSFLBAASLMSQSVYRHLVLLHGVCMAGDSTMVDFVHLGADIMYLRKRGHLVPASW
LOVVKOLAVALNVEDKGLPHGNVARKVLLAREGADGSPFIKISDPGVSFAVLSLE
MLTDRIPWVAPCLBEAOTLSLEADKMGFGATVWEVSGVTWPIISALDPAKLOFYED
ROOLPAKWTLEALLIQCMAYEPVORPSFRAVIRDLNLSLSDYELLSDPTPALAP
RDLGNGAGLYACQDPTIFERHLKYISQLGKNGFSGVSELCTFDPHAHTGALVAKQ
LOHSGPDQQRDFQREITQILKHSDFIVKRGVSGYGPRLVMEYHLPSCGLRDL
QHRARLDASRLLYSSQICKMGVYLRRCVRRDLAARNILVSEAHVKIADFLGAK
LPLDKDYVVRPQSPPIFWAPESLSDNIFSRQSDVMSFGVLYELFTYCDKSCSP
SAEFLRMGCERDVPAICRLLELBERGQRLPAPGAPAEVHELMKLCWAPSPQDRFSF
SALGPOLDMLWSGRCGCTHAFTAHPEGHHSLSFS"
BASE COUNT 746 a 1292 c 1168 g 858 t
ORIGIN
Alignment Scores:
Pred. No.: 3,82e-278 Length: 4064
Score: 4684.00 Matches: 905
Percent Similarity: 87.61% Conservative: 64
Best Local Similarity: 81.83% Mismatches: 129
Query Match: 79.93% Indels: 9
DB: Gaps: 5
US-09-397-967a-16 (1-1099) x HSU09607 (1-4064)
QY      1 MetAlaProProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20
Db      96 ATGCACCTCCAAAGTGAAGAGACGCCCTGATCCCTCAGCGTTTCATGACGCTCTGTC 155
QY      21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db      156 ACGGAGGCTGGTCCCTGTCATGTGCTGCTCCCGCTCGCGCCCGGGGGCCCCCAGCGC 215
QY      41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db      216 CTATCTTCTCTTTGGGACCACTTGGCTGAGACCTGTGCGTGAGGCTGCCAGGCC 275
QY      61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db      276 AGCGGCATCTGCTGTGTACCATCTCCTTTGCTCTGGCCACGGAGGACCTGCTCTGC 335
QY      81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db      336 TGGTTCCCCCGCAGCACATCTTCTCGTGAGGATGCCAGCACCCCAAGTCCTGCTGTAC 395
QY      101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db      396 AGGATTGCGCTTTTACTTCCCCCAATGTTGGCTGGAGAGTGCCACCGCTTCGGGCTA 455
QY      121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db      456 CGCAAGGATTGGCCAGTGTCTATCTTGCACCTGTCAGTCTCTGAGACACCTCTTTGCC 515
QY      141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160

```



516 CACCGAGTGCCTGGTGGGGGCGCTCCCGGTGGGCTCAGTCTCAAGGAGCAGGCT 575  
161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
576 GAGTCTCAGCCTGGCGGTGTGGACCTGGCCCGGATGGCGGAGAGCAGGCCCGAGCGG 635  
181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
636 CCGGAGAGCTGCTGAAGACTGCTCAGCTACAGGCTGCTACCCCAAGCCTGGCGAC 695  
201 ValIleGlnGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
696 CTGATCCAGGGCCTCAGTCTCGCGGAGGGCTATTCCGAGGACGGTCCGAGAGCC 755  
221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240  
756 CTGCC -GGCGTGGCGGCTGCCAGGACCGGCATCTGCTCATGGCCAAAGTACATCAT 814  
240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260  
815 GGACCTGGAGCGGTGGATCCAGCGGGGCCCGGACCTTCCAGCTGGGCTCCCTGG 874  
260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277  
875 GGCCTTGGTGGCCACGACGAGCGGCTGGGGCTGCTCGCGGTGGTGTGACGGCGCATCGC 934  
277 oTriSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296  
935 CTGGACCCAGGAGAACAGGAGGTCTCCAGGCCCTTCTGCGACTTTCAGAAATCGTAGA 994  
296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316  
995 CATTAGCATCAAGCAGGCCCCCGCGGTGGCCCGCGGAGACACCGCTTGGTCACTGT 1054  
316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336  
1055 TACCAGGACAGACACCAAGATTTTAGAGCGCGAGTTCCAGGGCTGCCCGAGGCTCTGTC 1114  
336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356  
1115 GTTCTGGCGCTCGTGGAGCGCTACTTCGGGTGACCCAGGACTCCCGACACTTCTTCTG 1174  
356 sLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376  
1175 CAAGGAGTGGCACCGCGCGAGCTCTGGAGGAAGTGGCGGAGCAGTGCACCGGCCCAT 1234  
376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396  
1235 CACTCTGGACTTTGGCATCAACAAGCTCAAGACTGGGGGCTCACGCTCCTGGCTCTATGT 1294  
396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416  
1295 TCTCGCGCGAGCCCCAGGACTTTGACAGCTTCTCCTCACTGTCTGTGTCAGAACCC 1354  
416 oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLe 436  
1355 CTTTGGTCTCTATTATAGSGTGCCTCATCGGCGCAGCCCCACAGGAACCTTCTTCT 1414  
436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTrpAsnSe 456  
1415 GTTTGGCTCAGCCGACCCACAGCAGTCTTCGAGAGCTCTCGGAACCTGCTGGGATGG 1474  
456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLy 476  
1475 GGGGCTGCAGCTAGATGGGTGGAGTACGCCCTCACTTCTGCTGTATCCCCAGACCCAA 1534  
476 sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496  
1535 AGAAAGTCCAACTGATCGTGTCCAGAGAGGTTCAGAGCCACCCACCATCATCTCTGTT 1594  
496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
1595 TCAGGCCCCCAATCCCAATACCAGTCTGAGTCAGATGACATTTTCAAGATCCCTGCTGACAG 1654

515 rLeuGlnTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535  
1655 CTTGAGTGGCATGAGAACCTGGGCCATGGGTCTTCCACCAAGATTTTCCGGGGGTGTCG 1714  
535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555  
1715 CCATGAGTGGTGGATGGGAGGCCCGAAGACAGAGAGGTGCTGCTCAAGGTCATGGATGC 1774  
555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575  
1775 CAAGCACCAAGAACTCATGGAGTCACTCTGGAAGCAGCAGCTTGTATGAGCAAGTGTGC 1834  
575 rTyrProHisLeuValLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595  
1835 GTACCGGCTATCTGCTGCTGCCAGGGCTGTCATGGTGGAGACAGCACCATGGTGCA 1894  
595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615  
1895 GGAAATTTGTACACCTGGGGGCCATAGACATGTATCTGGAAAAACGTGGCCACTTGGTGC 1954  
615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635  
1955 AGCCAGCTGGAAGCTGCAGGTGGTCAAACAGCTGGCTACGCCCTCAACTATCTGGAGA 2014  
635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655  
2015 CAAGGCTTGGCCCCATGCAATGTCTCTGCCGGAAGTGTCTCTGGGTCCGGAGAGGGGCGC 2074  
655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675  
2075 TGATGGAGCGCCCTTCATCAAGCTAGTGACCTGGGGTTCAGCCCGCTGTGTTAAG 2134  
675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695  
2135 CTTGAGATGTCTACCGCAGAGATCCCTGGGTGGCCCGCGAGTGTCTCTCGGAGGCGCA 2194  
695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715  
2195 GACCTTAGCTTGAAGCTGACAAAGTGGGGTCTCGGCGCCACGGTCTGGGAGTGTTAG 2254  
715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735  
2255 TGGCGTACCATGCCCCATCAGTCCCTGGATCTCTGCTAAGAAACTCCAATTTATGAGA 2314  
735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755  
2315 CCGGAGCAGCTGCGCGGCCCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGTGCAT 2374  
755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775  
2375 GGCCTATGAGCGGTGTCAGAGGCGCTCTCTCCGAGCGCTCATCTGTCACCTCATAGCCT 2434  
775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795  
2435 CATCTCTCAGACTATGAGTCTCTCAGACCCACACCTGTCCTGGTCCCTGGCACCTCGTA 2494  
795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815  
2495 TGGGCTGTGG---AATGGTGCCAGCTCTATGCTGCCAAGACCCACGATCTTCGAGGA 2551  
815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
2552 GAGACCTCAAGTACATCTCACAGCTGGGCAAGGCAACTTTTGGCAGCGTGGAGCTGTG 2611  
835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855  
2612 CCGCTATGACCCGTAGCCCAATACAGGTGGCTTGTGGCCCGTGAACAGCTGTCAGCA 2671  
855 sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875  
2672 CAGCGGCCAGACAGCAGAGGGGACTTTTTCAGCGGAGATTCAGATCCTCAAAGCACTGCA 2731

875 sSerAspPheIleValIleTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895  
 2732 CAGTGATTTCAITGTCAGATATCGTGTGTGACGATATGCGCGCGGCAGAGCTGCG 2791  
 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGl 915  
 2792 GCTGGTCAATGGAGTCTGCCAGCGGCTGCTTGGCGGACTTCCCTCAGCGGCGCCGCG 2851  
 915 y---LeuHisThrAspArgLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTy 934  
 2852 GCGCTCGATGCGCGCGCTCTCTCTATTCTCGCAGATCTGCAAGGGCATGAGAT 2911  
 934 rLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlus 954  
 2912 CTTGGGCTCCGCGCTGCGTGACCGGACCTGGCGCGCGCGAAACATCTCTGAGAG 2971  
 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974  
 2972 CGAGGCACACGTCAAGATCGTCACTTGGGCTAGCTAAGCTGCTGCGGCTTGACAAAGA 3031  
 974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLe 994  
 3032 CTACTACGTGTCCGGAGCGAGCGGACGCCATCTTCTGTATGCCCGCGAATCCCT 3091  
 994 uSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGl 1014  
 3092 CTGGGACACATCTTCTCTCGCAGTCAGACGCTTGGAGCTTGGGGTCTGCTGTACGA 3151  
 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGl 1034  
 3152 GCTCTTACCTACTGGGACAAAGCTGACGCCCTCGGCGAGTTCCTGGGATGATGGG 3211  
 1034 yProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgAr 1054  
 3212 ATGTGAGCGGATGTCCCGCCCTCTGCCCTCTTGGAACTGCTGGAGAGGGCCAGAG 3271  
 1054 gLeuProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTyrAl 1074  
 3272 GCTGGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3331  
 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094  
 3332 CCTAGCGCCACAGACCGCGCATCATTCAGCGCCCTCGGCGCGCTGCGCGCGCTGCTG 3391  
 1094 pArgGlyArgProGly 1099  
 3392 GAGCGGAGCGGGGG 3407

RESULT 7  
 AR270492  
 LOCUS 3620 bp DNA linear PAT 10-APR-2003  
 DEFINITION Sequence 1055 from patent US 6500938.  
 ACCESSION AR270492  
 VERSION AR270492.1 GI:29701726  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3620)  
 AUTHORS Au-Young,J. and Sellhamer,J.J.  
 TITLE Composition for the detection of signaling pathway gene expression  
 JOURNAL Patent: US 6500938-A 1055 31-DEC-2002;  
 FEATURES  
 Location/Qualifiers  
 1..3620  
 /organism="unknown"

BASE COUNT 670 a 1152 c 1076 g 722 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,27e-273 Length: 3620  
 Score: 4597.00 Matches: 895  
 Percent Similarity: 86.80% Conservative: 65  
 Best Local Similarity: 80.92% Mismatches: 123

Query Match: 78.45% Indels: 24  
 DB: 6 Gaps: 7  
 US-09-397-967A-16 (11-1099) x AR270492 (11-3620)  
 QY 1 MetAlaProSerSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
 DB 96 ATGGCACCTCCAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTCATGCAAGCTCTTGTTC 155  
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
 DB 156 ACGAGGCTGTGGCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 215  
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
 DB 216 CTATCTTCT 275  
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 DB 276 AGCGGATCTGCTGTGTACCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 335  
 QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 DB 336 TGGTTCCCGCGAGCCACATCT 395  
 QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
 DB 396 AGGATTCTGCTTTTCT 455  
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
 DB 456 CGAAGGATTGCGCCAGTGTCT 515  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
 DB 516 CACCGAGTGTGCTGT 575  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
 DB 576 GAGTGTCTCAGCTGCGCGT 635  
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
 DB 636 CCGGAGAGT 695  
 QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAla 220  
 DB 696 CTGATCCAGGCTGT 755  
 QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240  
 DB 756 CTGCG-CCGCGTGGCGCTGT 814  
 QY 240 uAspLeuArgGluLeuHisProAlaAlaThrGluThrPheArgValGlyLeuProGl 260  
 DB 815 GGACCTGTGAGCGGT 874  
 QY 260 YAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277  
 DB 875 GGCCTTGT 934  
 QY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296  
 DB 935 CTGGACCCAGGAGAACAGGAGGTCTCTCCAGCGCTTCTGCGACTTCTCCAGAAATCTGTAGA 994  
 QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316  
 DB 995 CATTAGCATCAAGCAGGCGCGGT 1054  
 QY 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336  
 DB 1055 TACGAGACAGACAAACAGATTTTAGAGGCGGAGTTCACAGGGCTGCCCGAGGCTCTGTCT 1114



336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356  
1115 GTTCGTGGCGCTCGTGGACGGCTACTTCGGCTGACCCAGGACTCCACGACTTCTTCG 1174  
356 slysgluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIl 376  
1175 CAAGAGAGTGGCACCGCGAGCGTGTGGAGAAAGTGGCGAGCAGTGCACGGGCCCAT 1234  
376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIl 396  
1235 CACTCTGGACTTTGGCATCAACAGCTCAAGACTGGGGGCTCAGCTCCGGCTCCTATGT 1294  
396 eLeuArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416  
1295 TCTCGCGCGACGCCCGCCAGGACTTTGACAGCTTCCTCTCACTGTGTGTCCAGAACCC 1354  
416 oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerIe 436  
1355 CTTTGGTCTGTATTATAAGGGCTGGCTCATCCGGCGCAGCCCGCACGGAACCTTCCTCT 1414  
436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTrpAsnSe 456  
1415 GGTTCGGCTCAGCCGACCCACAGCAGTCTTCGAGAGCTTCCTGGCAACCTGTGGATGG 1474  
456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProly 476  
1475 GGGGCTGCACGTAGATGGGTGGCAGTACCCTCACTTCCTCTGTATATCCCGACGCCAA 1534  
476 sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496  
1535 AGAAAGTCCAACTCGATGGTGGTGCAGAGAGGTACAGCCCGCACCATCATCTCTGGT 1594  
496 sSerPro---SerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
1595 TCAGCCCCAATCCCAATACCAGCTGAGTCAGATGACATTTTCAAGATCCCTGCTGACAG 1654  
515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535  
1655 CTTGAGTGGCATGAGAACCTGGGCCATGGGTCTTCACCAAGATTTACCGGGGGTGTGC 1714  
535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555  
1715 CCATGAGTGTGTGATGGGAGCGCCGAAAGACAGAGGTGTCTGCTGAAGGTCAATGGATGC 1774  
555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575  
1775 CAAGCACAGAAGTGCATGGAGTCAATCTCTGGAAGCAGGAGCTTGATGAGCAAGTGTTC 1834  
575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595  
1835 GTACCGGCATCTCGTGTCTCCACGGCGTGTGCATGGCTGGAGACAGCACCATGGTGCA 1894  
595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615  
1895 GGAATTTGTACACCTGGGGGCGCATGACATATATCTGGAACACGTGGCCACCTGGTGCC 1954  
615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635  
1955 AGCCAGCTGGAAAGCTGCAAGTGGTCAACAGCTGGCCTACGCCCTCAACTATCTGGAGGA 2014  
635 pLyGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655  
2015 CAAAGGCTTGGCCCATGCAATGTCTCTGCCCGGAAGGTGCTCCTGGCTCGGGAGGGGGC 2074  
655 yAspGlyAsnProPropheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675  
2075 TGATGGAGCCCGCCCTTCATCAAGCTGAGTCACCTGGGGTCAGCCCGCGCTGTCTTAAG 2134  
675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695  
2135 CTTGAGATGCTCAACGACAGGATCCCTGGGTGGCCCCCGAGTCTCCGGGAGGGCGCA 2194  
695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715

2195 GACACTTAGCTTGAAGCTGACAACTGGGCTTCGGGCGCACGGCTCGGAAGTGTITAG 2254  
715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735  
2255 TGGCGTCAACCTAGCTGCGCTGAGTCTCTGCTAAGAACTCCAATTTTATGAGGA 2314  
735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysSe 755  
2315 CCGGACAGAGCTTGGCGGCCCAAGTGGACAGCTGGCCCTGCTGATTCAACAGTGCAT 2374  
755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775  
2375 GSCCTATGAGCGGTCAGAGGCCCTCTTCGAGCCGCTCATTCGTGACCTCAATAGCCT 2434  
775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795  
2435 CATCTCTTCAGACTATGAGTCTCTCAGACCCACAGCTGGTGGCCCTGGCACCTCGTA 2494  
795 pGlnGlyCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815  
2495 TGGGCTGTGG---AATGGTGCCAGCTCTATGCCCTGCCAAGACCCACGATCTTCGAGGA 2551  
815 uArgHisLeuLysTyrIleSerLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTG 2611  
835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855  
2612 CGGCTATGACCGCTAGCGGACAAATACAGGTGCCCTGGTGGCGCTGAAACAGCTGCAGCA 2671  
855 sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaIleHi 875  
2672 CAGCGGGCAGACACAGCAGAGGAGCTTTACGGGGAGATTTCAGATCCTCAAGACTGCA 2731  
875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895  
2732 CAGTGATTTTCAATGTCAGTATCGTGGTGTCACTATGGCCGGCGCCGACAGCTGCG 2791  
895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGl 915  
2792 GCTGTCTATGAGTACCTGCCCGCGGTGTGTGGCGACTTCCTGCGGCGCACCGCGC 2851  
915 y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyr 934  
2852 GCGCTCGATGCCCGCCCTCTCTCTATTCCTGCGAGATCTGCAAGGGCATCGAGTA 2911  
934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954  
2912 CTTGGGCTCCCGCGCTCGTGCACCGGACCTGGCCGCCCGCAACATCTCTCGTGAGAG 2971  
954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974  
2972 CGAGCACACGCTCAAGATCGCTGACTTCGGCTAGCTAAGCTGTCTGCGCTTGACAAAGA 3031  
974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerIe 994  
3032 CTACTACGTGTGTCGCGAGCGCAGCGCCAGAGCCCAATTTCTGGTATGCCCGCAATCCT 3091  
994 uSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGl 1014  
3092 CTGGACACATCTCTCTCGCCAGTCAGAGCTCTGGAGCTTCGGGGTCTGCTGTAGCA 3151  
1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGl 1034  
3152 GCTCTTACCTACTCGGCAAAAGCTGACGCCCTCGGCCGAGTTCCTCGCGATGATGGG 3211  
1034 yProGluArgGluGlyProProLeuCysArgLeuLeuLeuAlaGluGlyArgAr 1054  
3212 ATGTGAGCGGATGTCCCGCCCTCTGCGGCTCTTGGAACTGTGGAGGAGGGCCAGAG 3271  
1054 gLeuProProProThrCysProThrGluValGlnGlnLeuMetGlnLeuCysTrpAl 1074



815 GGACCTGGAGCGGCTGGATCCACCGCGGCGCCGAGACCTTCCACGTGGGCTCCCTGG 874  
260 yala-----GlnGluGluProGlyLeuArgValAlaGlyAspAsnGlyIlePr 277  
875 GGCCTTGGTGCCAGCAGCGGCTGGGCTGCTCCGCGTGGCTGGTGAGCGGCGATCGC 934  
277 ofrpSerSerAnApp--GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296  
935 CTGGACCCAGGAGAGACAGAGGTCCTCCAGCCCTTCTCGGACTTTCAGAAATCGTAGA 994  
296 pValSerIleAnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316  
995 CATTAGCATCAACAGAGCGCCCGCGCTTGGCGCGCGGAGAGCACCGCGCTGCTCACTGT 1054  
316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336  
1055 TACCAGGACAGAACACAGATTTAGAGCGCCGAGTTCCCGAGGCTGCCCGAGGCTCTGTC 1114  
336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisIleTyrPheCy 356  
1115 GTTCGTGGCGCTCGTGAGCGGTACTTCCGGCTGACCACGGACTCCAGCACTTCTCTCG 1174  
356 sIysGluValAlaProProArgLeuLeuGluAlaAspValCysHisGlyProIle 376  
1175 CAAGGAGGTGGCACCGCGAGGCTGCTGGAGGAAGTGGCGGAGCAGTGCCACGGGCCCCAT 1234  
376 eThrLeuAspPheAlaIleHisIleLeuLysAlaAlaGlySerLeuProGlyTyrIle 396  
1235 CACTCTGGACTTGGCATCAACAGCTCAAGACTGGGGGCTCAGGTCTCGGCTCTGATGT 1294  
396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416  
1295 TCTCGCGCGACGCCCGAGGACTTTGACAGCTTCTCTCTCACTGTGTGTCCAGAACCC 1354  
416 oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerIle 436  
1355 CTTGGTCTCTGATTATAAGGCTGCTCATCGGCGGCGAGCCCGACAGGAACCTTCTCTCT 1414  
436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSe 456  
1415 GGTGGCTTCAGCCAGCCCGACAGGAGCTTTCGAGAGCTCTCGGCAACCTGCTGGATGG 1474  
456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProly 476  
1475 GGGGCTGCACGTAGATGGGTGGAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1534  
476 sGluLysSerAnLeuIleValValArgGlyCysAsnProAlaProAlaProGlyCy 496  
1535 AGAAAGTCCAACTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1594  
496 sSerPro--SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
1595 TCAGGCCCCAATCCCAATACCACTGAGTCAGATGACATTTCAAGATCCCTGCTGAGAC 1654  
515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrIlePheArgGlySerAr 535  
1655 CTTGAGTGGCATGAGAACCTGGGCGCATGGTCTTTCACAGATTACCGGGGCTGTCG 1714  
535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555  
1715 CCATGAGTGGTGGATGGGAGCGCCGAAAGACAGAGGTGCTGCTGAAGTTCATGGATGC 1774  
555 rArgHisArgAnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575  
1775 CAAGCACAGAACTGCATGGAGTCAATCTGGAAGCAGCGAGCTTGTGAGCAAGTGTTC 1834  
575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595  
1835 GTACCGGCTCTCTGCTGCTCCAGCGGTGTGATGGCTGGAGACAGCACCATGGTGA 1894  
595 nGluPheValTyrLeuGlyAlaAlaAspMetTyrLeuArgLysArgGlyHisLeuValSe 615  
1895 GGAAATTGTACACCTGGGCGCATAGACATATCTGCGAAACCTGGGCCACTGGTGGCC 1954

QY 615 rAlaSerTrpLysLeuGlnValThrIysGlnLeuAlaTyrAlaLeuAsnTyrIleuGluAs 635  
Db 1955 AGCCAGCTGGAGCTGACGCTGCTCAACAGCTGGCTTACGCCCTCAACTATCTGGAGA 2014  
QY 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655  
Db 2015 CAAGGCTTGGCCCATGGCAATGCTCTGCCCGGAAGTGTCTCTGGCTCGGAGGGGGC 2074  
QY 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675  
Db 2075 TGATGGAGCGCCCGCTTTCATCAAGCTGAGTCACTGGGGTCAAGCCCGCTGTGTAAAG 2134  
QY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695  
Db 2135 CTTGAGAGTGTCAACGACAGGATCCCTGGGTGGCCCGGAGTGTCTCCGGAGGCGCA 2194  
QY 695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715  
Db 2195 GACACTTAGCTTGGAGGCTGACAGTGGGGCTTCCGCCCGCAGGCTCTGGGAAGTGTTAG 2254  
QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735  
Db 2255 TGGCGTCAACATGCTCAGTGCCTGGATCTCTGCTAAGAACTCCAATTTATGAGA 2314  
QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755  
Db 2315 CCGGAGCAGGCTGCGCGGCCCCCAAGTGGACAGAGTGGCCCTGCTGATTTCAACAGTGCAT 2374  
QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775  
Db 2375 GGCCTATGAGCGGTCCAGAGGCCCTCTCTCGAGCGCTCATTCGTGACCTCAATAGCT 2434  
QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795  
Db 2435 CATCTCTTCAAGTATGAGTCTCTCTCAGACCCACACCTGCTGCTGCTGCTGCTGCTG 2494  
QY 795 pGluLeuCysValAlaAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGl 815  
Db 2495 TGGGCTGTGG--AATGGTGGCCAGCTCTATGCTTGCACAGACCCACAGATCTTCGAGA 2551  
QY 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
Db 2552 GAGACACCTCAAGTACATCTCAGCTGGGCGAAGGGCAACTTTTGGCAGCGTGGAGCTGTG 2611  
QY 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855  
Db 2612 CCGCTATGACCGCTAGCGCAATACAGGTGCCCTGCTGGCGCTGCTGCTGCTGCTGCTGCTG 2671  
QY 855 sSerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875  
Db 2672 CAGCGGCGCAGACCGCAGAGGAGCTTTCAAGCGGAGATTCAGATCCTCAAGACACTGCA 2731  
QY 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895  
Db 2732 CAGTGATTTTCATTGTCAAGTATCGTGGTGTGAGCTATGGCCCGGCGCCGAGCGCTGCG 2791  
QY 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGl 915  
Db 2792 GCTGCTCATGAGTACCTGCTGCCAGCGGTGCTTGGCGACTTCTCTGCGAGCGCACCGCGC 2851  
QY 915 y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTy 934  
Db 2852 GCGCTCGATGCCAGCGCCCTCTCTCTATTCCTCGCAGATCTGCAAGGGCATGGAGTA 2911  
QY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954  
Db 2912 CTTGGGCTCCCGCGCTGCTGCGCCGACCTGGCCCGCCCGCAACATCTCTCTGGAGAG 2971  
QY 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974  
Db 2972 CGAGGCACACGTCAAGATCGTACTTGGCGCTAGCTAAGCTGCTGCTGCTGCTGCTGCTGCTG 3031

QY 974 pYrVrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLe 994  
Db 3032 CTACTAGTGGTCCGGAGCCAGCCAGCCCAATTTCTGTATGCCCCCAATCCCT 3091  
QY 994 uSerAspAenIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGl 1014  
Db 3092 CTCCGACACATCTTCTCCGAGTCAGACGTCGTGGAGCTCGGGTCTGCTGTACGA 3151  
QY 1014 uLeuPheThrTyrCysAspIlySerCysSerProSerAlaGluPheLeuArgMetMetGl 1034  
Db 3152 GCTCTTACACTACTGACAAAAGCTGACGCCCTCGCCGAGTTCCTCGGAGTGGG 3211  
QY 1034 YProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgAr 1054  
Db 3212 ATGTGAGGGAGTCTCCCGCCCTCTGCCGCCCTCTGGAACCTGCGAGGAGGCCAGAG 3271  
QY 1054 gLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074  
Db 3272 GCTGCCGGCGCTCTGCTGCCCTGCTGAGGTG----- 3305  
QY 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094  
Db 3306 -----AGCCCGAGGGCTAGCTCAGTTCCAGTCTGTAGAT-----TG 3346  
QY 1094 pArgGlyArgProGly 1099  
Db 3347 GCGCGGGTCTCGGC 3362

AC073700 218208 bp DNA linear HTG 29-JUN-2000  
LOCUS Mus musculus clone RP23-162012, WORKING DRAFT SEQUENCE, 16  
DEFINITION unordered pieces.

ACCESSION AC073700  
VERSION AC073700.1 GI:8810317  
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 218208)  
TITLE DOE Joint Genome Institute.  
JOURNAL Sequencing of Mouse  
REFERENCE 2 (bases 1 to 218208)  
TITLE DOE Joint Genome Institute.  
JOURNAL Direct Submission  
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov  
-----  
Project Information  
Center Project Name: 1799428  
Center clone name: RPCI-23\_162012  
-----

Summary Statistics  
Consensus quality: 204328 bases at least Q40  
Consensus quality: 210800 bases at least Q30  
Consensus quality: 212488 bases at least Q20  
Estimated insert size: 222000; agarose-fp estimation  
Estimated insert size: 216708; sum-of-contigs estimation  
Quality coverage: 9.37 in Q20 bases; agarose-fp estimation  
Quality coverage: 9.6 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a "working draft" sequence. It currently  
\* consists of 16 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will

\* be preserved.  
\* 1 1092: contig of 1092 bp in length  
\* 1093: gap of unknown length  
\* 1193: contig of 1373 bp in length  
\* 2565: gap of unknown length  
\* 2656: gap of unknown length  
\* 2666: contig of 1046 bp in length  
\* 3711: gap of unknown length  
\* 3812: contig of 1744 bp in length  
\* 5555: gap of unknown length  
\* 5556: gap of 3359 bp in length  
\* 9014: gap of unknown length  
\* 9114: contig of 3410 bp in length  
\* 9115: contig of 3410 bp in length  
\* 12525: gap of unknown length  
\* 12525: contig of 5691 bp in length  
\* 18315: gap of unknown length  
\* 18316: contig of 6884 bp in length  
\* 18416: gap of unknown length  
\* 25299: gap of unknown length  
\* 25300: gap of unknown length  
\* 25400: contig of 9459 bp in length  
\* 34858: gap of unknown length  
\* 34859: contig of 13153 bp in length  
\* 48111: gap of unknown length  
\* 48112: contig of 11464 bp in length  
\* 59675: gap of unknown length  
\* 59676: contig of 19464 bp in length  
\* 59775: gap of unknown length  
\* 59776: contig of 23420 bp in length  
\* 79240: gap of unknown length  
\* 79340: contig of 27114 bp in length  
\* 102759: gap of unknown length  
\* 102760: contig of 27114 bp in length  
\* 102859: gap of unknown length  
\* 102860: contig of 34115 bp in length  
\* 129773: gap of unknown length  
\* 130074: contig of 34115 bp in length  
\* 130074: gap of unknown length  
\* 164189: contig of 53920 bp in length.  
\* 164289: contig of 53920 bp in length.  
\* 218208: contig of 53920 bp in length.

FEATURES  
Location/Qualifiers  
1..218208  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="RP23-162012"  
/clone\_lib="RPCI mouse BAC library 23"  
BASE COUNT 51066 a 56856 c 56653 g 52132 t 1501 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.2e-239 Length: 218208  
Score: 4091.50 Matches: 1072  
Percent Similarity: 32.75% Conservative: 2  
Best Local Similarity: 32.69% Mismatches: 24  
Query Match: 69.82% Indels: 2188  
DB: 2 Gaps: 24

US-09-397-967A-16 (1-1099) x AC073700 (1-218208)  
QY 1 MetaAlaProSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
Db 110153 ATGGACCTCCAGTGAGGACACCTCTGATCCCTCAGGCTCTTCAGCTCTATCC 110212  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
Db 110213 TCAGAGCAGGAGCCCTGCATGTCTCTCTCCCGGGAGCTGGCCCGCAGCGA 110272  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaIlySala 60  
Db 110273 TTGTCAATCTCTTTGGGGACTACTTGGCTGAGGATTTATGTGTCGAGCTGCCAAGGCC 110332  
QY 61 CysGlyIle----- 63  
Db 110333 TGTGTGA-GTGGGTCCCTAGATGACCATGCTATGTAACATCTATTGAAAGG 110391  
QY 63 ----- 63  
Db 110392 GGATGCTGGGGTTTTGAGGGGTGTGGAATTTTGTAAAGGAGGATCATGGATCAGCCCG 110451  
QY 64 -----LeuProValTyrHis 68

Db	110452	TGTAAGTGGCTGGTGAACAGTATGTCGCCATCCCTTCAGGCATCCTCGCTGTTATCAT	110511	Db	111531	CTGGAGCGGTACATCAGCGGCCACACCGAGACCTTCGTGTGGGCTCCGGGGGCC	111590
QY	69	SerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProProSerHisIlePhe	88	QY	262	GlnGluGluProGlyLeuArgValAlaGlyAspAsnGlyIleProTrpSerSerAsn	281
Db	110512	TCCTTTTCGCTCGCCACTGAGGACTCTCTGTGTGGTTTCCCCCAAGCCATCTTC	110571	Db	111591	CAGGAGAGCCGGGGTCTTCGTGTGGGGGGACAACGCGCATCTCCTGAGCTCGGG	111650
QY	89	CysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg	103	QY	282	AspGlu	283
Db	110572	TGCATAGAGGACGTGGACACTCAAGTCTTGGTCTACAGGCTACGGTAGGAAGTGAACCCC	110631	Db	111651	GACCA-GGAGGTGCTGGGGCTGGGCCCTTGAGAAAGCTGGGTCCGCGGGGAGCGCTGGCGGT	111709
QY	103	-----	103	QY	283	-----	283
Db	110632	TTGAATTTATCTTCAGGGTAGTATTCTATTTCAGTCTATAGTAGGCTCTTGAAACAGA	110691	Db	111710	TAGTAAAGAAGCCCGTCCGAACAATCAGGAAGTGGCTAGTCCGGCGGGGACTGATTGCA	111769
QY	103	-----	103	QY	283	-----	283
Db	110692	GAGGTGGGCTGTTTCTGTAGGCCACACTGCACCTTCAGTGACCCAACTTGGGCTCTGGG	110751	Db	111770	CTACGGGGCAGGGGGCGGGGTCAATTAAGGTGGGTCCGGGGATGGAGCTACCCAGGA	111829
QY	104	-----PheTyrPheProAspTrp	109	QY	283	-----	283
Db	110752	AAGCCAGAGCTGAGTGGGCCACCCCATTCGAAACCCACAGCTTTTATTTCCCTGACTGG	110811	Db	111830	CTGGACCTGGGGCGGTGACTAGAAAGGTGAGACTAGATAGGGGGCGGGACGGACAGAAG	111889
QY	110	PheGlyLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeu	129	QY	283	-----	283
Db	110812	TTTGGGCTGGACATGTACCGCTTTGGGCTGCCAAGATTGACAGTGCATCCTT	110871	Db	111890	GCGATAGTATGGATTGGATCGGTGGGTGTGACCCGACAGGCTGTCTCAGTCCCCCTAAC	111949
QY	130	AspLeuHisValLeuGluHisLeuPheAla	139	QY	283	-----	283
Db	110872	GACTTACAGTTTTAGAACATCTCTTTGCTCAGTTGGTATTCTTCGGTGGGAAGGCTC	110931	Db	111950	CTCGTCCCCGACCCCATCCACCCCTCCCGCTAAAGTGGAGGTGGAGAGGAGAAG	112009
QY	139	-----	139	QY	283	-----	283
Db	110932	AACCCAGGATGGGGACCCCAAGTTAACAAATCCAGCACCCAGACCGAACTCACCT	110991	Db	112010	AAGGGAGGCTGTGTGGGTTCAGGTCTGTAATCTCTTGCTTCTGCTGTCTCTTTCT	112069
QY	140	-----GlnHisArgSerAspLeuValSerGlyArgLeuProVa	152	QY	283	-----	283
Db	110992	AACCAACCTTCTGTGGCTGACCCAGCAGCCGAGTACCTGGTGGGCGCTCCCGGT	111051	Db	112070	CTCTTTTTTTAAATGGTTTTTGGAGACACAGGTTTTCTCTGTATAGTCTTGCTGTCTCGGA	112129
QY	152	LgylLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaG	172	QY	283	-----	283
Db	111052	GGGCTTAGCATGAAGGAGCAGGAGAGTCTCTGAGCCTGGCGCTGTGACTTGGCCCA	111111	Db	112130	ACTCATTTTGTAGACCAAGCTGGCTGGAACTCAGAAATCCGCTCTGCTCTCGCTCGCA	112189
QY	172	nMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrVal	188	QY	283	-----	283
Db	111112	GATGCTGTGAGAGGCCCCAGCGCCAGAGAGTGTCTGAAGACGGTCAGGTAGGGGCC	111171	Db	112190	GTGCTGTGATTAAGGCGTGCGCCACTACACCCAGCTACTGCTGTCTCTTCTCTCTCAG	112249
QY	188	-----	188	QY	284	LeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaPro	303
Db	111172	TCGGGCCACCCCTTCGTGTCTTCTTTGGCTGGGATCTAGGGCCCTATCTGTTTCCCTT	111231	Db	112250	CTTTTCCAGACCTTCTGTGACTTTCCGGAAATCGTGGATGTGAGCATCAAGCAGGCCCA	112309
QY	188	-----	188	QY	304	ArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIle	323
Db	111232	ACTGGAACACATACCTCTCTCCAACTGTGTGAGCCACCCCTCTCCGTGAGCCTGCCCTG	111291	Db	112310	CGTGTGGTCCGCGAGGGGAACACCGGCTGGTCACTCCGTCAACAGGATGACCGCCACATC	112369
QY	188	-----	188	QY	324	Leu	324
Db	111292	CGAAAGGTCCCTATGTCTGTCTCGTGTCTCTTGGCTGGGATCTAGGGCCCTATCTG	111351	Db	112370	CTGGTGGTCTTTGGCACCCGTTATTTCTGGTGTGATCTCTGGAAAGGGGTGGGTTCCGTT	112429
QY	189	-----SerTyrLysAlaCysLeuProProSerLeuArgAspValI	202	QY	325	-----GluAlaGluPheProGlyLeuProGluAlaLeuSe	336
Db	111352	TCGGGTGGCCCCCCCCACAGTTACAAAGCTGTCTGCCGCCACAGCTCGCGCATGTA	111411	Db	112430	GTTCGTGAAGAGTTCTCTCTGCGCAGGAAGCGGAGTTTCCGGGCTGCTTGAGGCCCTGTC	112489
QY	202	IleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAlaLeuL	222	QY	336	rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy	356
Db	111412	TCAGGGCCAGAACTTCGTGACACGACGCGCATCCGAGGACCGTGTCTTGGCGCTGC	111471	Db	112490	CTTCGTGGCCCTTGTGATGGGTACTTCGCGCTGATCTGCGACTCCAGGCATTAATTTCTG	112549
QY	222	euProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeuAsp	241	QY	356	slySGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIle	376
Db	111472	G-CCGTGTGGTGCCTTGCCAGGCCACCGCTACGGCTCATGGCCAAGTATATTTGGAC	111530	Db	112550	CAAGAGGTGGCGCGCGCGCTGCTGGAGGAGCGGAGCTGTGCCATGGAACCAT	112609
QY	242	LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAla	261	QY	376	eThr	377
		-----		Db	112610	CACGTAGGAGCTAGTTGGGGTTCACAGTGGGGATGGGGCTGGGCAGCGGAGGACTGTCTGG	112669



Db	114828	CCCTGGCCTCTCTCCACCCCTGCTCTACCTGCTCATGCTCTGTATACCTTAAGATCATCAC	114887	Db	115907	ACAGGCTAGCCTCACCCCATATATACCCCGTGTCTGTGACGCCACACGTGTGTACACAGGSCCT	115966
QY	591	-----	591	QY	678	-----	678
Db	114888	CCGGTAGCACCACCACTCATATCCAGGCCACACCCGGGCACACCCGATTCATTTCCACC	114947	Db	115967	GCACTGGCAAAAGGGGACCCCTTTGTGCAAAAGGTGACCATGACAGTTCTCACTCCCACGTTC	116026
QY	591	-----	591	QY	679	-----	696
Db	114948	TTGCCAGCTAATTGCTATGCAACATTCCTCCACCCACCCCGTGTGCCAACCCCTCCTGA	115007	Db	116027	TCAGTGTCTACCGACAGAAATACCCCTGGGTGGCCCCCGAATGCTCCAGAGGCTCAGAC	116086
QY	591	-----	591	QY	696	rLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnAr	716
Db	115008	ATCTTGCTGTATCTATTTCAGTTGAGTGTATAGAAACAGAACTCCAGCCTGTGCAAGA	115067	Db	116087	ACTCGGCTTGGAGGCTGACAAAGTGGGGCTTTGGAGCCACCACTGGAGGTGTTCACGG	116146
QY	591	-----	591	QY	716	gGlyProAlaHisIleThrSerLeuGluProAla-Lys	728
Db	115068	GCAGAGGAAGATGGGGGAGGGCCCTGTGAGTACTGTCTACCTGATCACTCAACCCACCC	115127	Db	116147	GGGACCCGCCCATCATCCCTCGCTGGAGCCCGCAAAAGTCAGAACCGTGGTCTCGGCCT	116206
QY	592	-----	604	QY	728	-----	728
Db	115128	ACTCTCCCTCTGCCAGGCATCATGTGTCAGAAATTTGTGTATCTAGAGCAATTGAC	115187	Db	116207	TCCTGGSCCGGTGAGGGGAGGGGATATCGTGAGGGTAGGGTTAAATCCGAGCTGAGCA	116266
QY	605	MetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLys	624	QY	729	-----	738
Db	115188	ATGTACTCGCAAGCGTGGCCACCTGTGTGAGCCAGCTGGAAACTGCAGGTGACCAAG	115247	Db	116267	TCGCAACACCGAAGTGAACCTCGTCCCTTAGAAGCTGAAGTTCATATGAGGACCCAGGAC	116326
QY	625	GlnLeuAlaTyrAlaLeuAsnTyr	632	QY	738	lnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrA	758
Db	115248	CAGCTGGCATAAGCCCTTAACCTACTTGGTGAGTGTCCTGTGTCCATTAAGGAGGTTCT	115307	Db	116327	AGCTGCCCTCTCAATGACAGAACTGGCGGACTTATCACACAGTGCATGGCCCTATG	116386
QY	632	-----	632	QY	758	spProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAenGlyLeuIleThrS	778
Db	115308	CTATGGATGGGAAATGAGAGTAGGTTCTGTAGTTTGGATAGAGGTTGGTAAGACACA	115367	Db	116387	ATCCTGGCCGGCGCCCTCTCTCCGAGCTATCTCAGAGCTTCAACGCGCTCAATTACAT	116446
QY	632	-----	632	QY	778	er	778
Db	115368	AGCAAAACGTTCCAGATGACGTGACGAGCTCAGAGGGCTGTAAAGTCAAGGAGGACA	115427	Db	116447	CAGGTCCGGAGAGGGGGGGGCGGAGGGTCCCTGAGCTGGGTCCAGGAA	116506
QY	632	-----	632	QY	778	-----	778
Db	115428	CTGTGAACCCATGGGTGTCTAGGGAAGACGTGAGTGAGGAGGCCAGGCTGAGGGGTGGG	115487	Db	116507	GTTTGTGTTTCTAGCAGTCAAGCTGCTGCAACAGCTGCTGCAACAAACCTAAGACC	116566
QY	632	-----	632	QY	778	-----	778
Db	115488	CGGGGTGTAGAGTGGAAACTGGTGAGAGAGGGGGTGTCTTGAATGGCGGATGTGGG	115547	Db	116567	GATCTGATTTGTTTGCAGTGTAGCCAGGTGGTGTTTTACAATGTAGCCCAAACTAG	116626
QY	632	-----	632	QY	778	-----	778
Db	115548	CTTTAACTGTCCCTAAAGTCGTGGGAAGGTAGAGAAAGGGTGTGGTGCTTAAGGATGG	115607	Db	116627	CATCTCAGGCTGGATTTCTCCCTGCTTCACCATCATGGTGGTGGGATTTGGTTTATTT	116686
QY	632	-----	632	QY	778	-----	778
Db	115608	TATGTCAGTGAAGGGGAGGTAGGGTTACAAGACATGACTTTCAGCGCCCTACAGTGT	115667	Db	116687	TGTTATTTCTAAATTTTGTTCGACAAAGATGTTCTCATCTGAGGGTATCCAAGTCAAGACT	116746
QY	632	-----	632	QY	778	-----	778
Db	115668	TGCCAAAGCCTGTGTGTGTAGGCAGCATCAGTAGGCTTTTGAAGATGCAGGAGTTT	115727	Db	116747	GGGTCATGATGGCTGTGCTCTCGGGTCCCTTTTGTCTGTAGCCGTTGGGGTCCCA	116806
QY	633	-----	640	QY	778	-----	778
Db	115728	GCTCAACATACTTCCCTGTTCCCTGTAACTTAGGAGGACAAAGGCTTCTCTCA	115787	Db	116807	CAGAGTGTGCTTTTGCCTGGACAGTGGTCTCTCATTTCCCATTAAGTACTTGTGGGG	116866
QY	640	sGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyLysAspGlyAsnProPr	660	QY	778	-----	778
Db	115788	CGGCACGCTCAGACCGAAGGTGCTCTGCTGTGAGGGGGGTGATGGGAATCCACC	115847	Db	116867	TCTCATAACGCACACAGAGAGCAAGGTTTCAGGTGGAGTTTACACCAATTTCAGTTTGG	116926
QY	660	oPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMet	678	QY	779	-----	781
Db	115848	TTTCAATTAAGCTAGTATCTCTGGTGTAGTCCCACTGTGCTGAGCTGGAAAG-TGAGT	115906	Db	116927	ATTGAGGCAGAGACTTACCACTTACCCCACTTCCTCAGTGATCCCGCCCTAGATTACG	116986
QY	678	-----	678	QY	781	luLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAlaG	801
Db				Db	116987	AGTCTCTCAGACCCCAACACCTGGCATCCCGATCCCGAGTCTTCGAGATGAGCTGTGC	117043



```
QY 801 l yAlaGlnIleuTy rAlaCysGlnAspProAlaIlePheGluGluArqHisLeuLysTy rI 821
Db 117044 GC C C C C A G C T C T A T G C C T G C A G A C C C C C C A T A T T C G A G G A G A C A C C T T A A G T A C A 117103
QY 821 l eSerLeuLeuGlyLys- 826
Db 117104 T C T C T T T G C T G G C A A G G T C A G T G G G C G G C A T G T G G G G A G N A C G T G G T G G T G A T 117163
QY 827 -----GlyAsnPheGlySerValGluLeu 834
Db 117164 G G G T C A G G T G G A C A C T C C C T C T C A T C C C C A C A G G C G A A C T T T G C A G C G T G G A G C T G 117223
QY 835 CysArgTy rAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln 854
Db 117224 T G C C G C T A T A C C C C C T G G G G A C A A T A C G G A C C C T G T G G C A G T G A A C A C G T A C A G 117283
QY 855 HisSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnIleLeuLysAlaLeu 874
Db 117284 C A C A G C G G C C A G A C C A G A G A G G A C T T C C A G C G G A G A T T C A G A T C C T T A A G G C T C T G 117343
QY 875 HisSerAspPheIleValLysTy rArgGlyValSerTy rGlyPro- 889
Db 117344 C A C A G C A C T T C A T C G T C A A G T A C C G G G A G T C A G T A T G G C C A G G T G A G C G C A G A C A G 117403
QY 889 ----- 889
Db 117404 C A T C T G G G A A C G G G T T C G A G T C C G C T T C T A C C C T T T C C A G T A G G C C C T G T G A C A 117463
QY 889 ----- 889
Db 117464 A G G T C G T T A A C T C C C A T C C C A G C T T C T A T A G T G G A G G T T G A A C G A A T C A A T A C A C 117523
QY 890 -----GlyArgGlnSerLeuArgLeu 896
Db 117524 C C A G T A G A A G G T G C T C G A G A G C T A C T G C G T G T T C G C A G G T C C C A G A G C C T C G G T T G 117583
QY 897 ValMetGluTy rLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly- 915
Db 117584 G T A A T G A G T A C C T G C C C A G C G G C T G C T G C A G A C T T C T T G C A G C G C C A T C G C G C G C G 117643
QY 916 LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys- 930
Db 117644 C T G C A C A C C G A C C G C T A C T G T T C G T T G G C A G A T C T C A A G G T C G A G A T C G C G G 117703
QY 930 ----- 930
Db 117704 A G A T C G G A T G G G G T C T C T G G G G T C C G G A T C C G G C C C A C G T G C T A T A T A T G G C T C T C T 117763
QY 931 -----GlyMetGluTy rLeuGlyValArgArgCysValHisArgAspLeuAlaAlaArgAs 949
Db 117764 G C A G G C A T G A G A T A C C T G G G T G C C G C G C T G C G T A C A C G T G A C C T G C T G C G C G C A A 117823
QY 949 nIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLe 969
Db 117824 C A T C T T G T G G A G A C G A G G C T C A T G T A A G A T C G C G A C T T C G G C C T C G C T A A G C T G C T 117883
QY 969 uProLeuGlyLysAspTy rTy rValValArgGluProGlyGlnSerProIlePheTrp- 988
Db 117884 G C C C C T G G A A A G G A C T A C T A G T G T C C G A G C C T G G C A A G C C C A T C T T T T T G T A 117943
QY 988 ----- 988
Db 117944 C A A G A C C A A G A C T C C A G A G G T C A C G C C C A C T G A G C C C C A C T C C C G C A C T C T A C A G C C T T 118003
QY 988 ----- 988
Db 118004 A A A C C G G T C C A G A C C A G C T C C A C C T C A G T A G C T T T G C C C T G C C A A G C C C T T G C C A C A 118063
QY 988 ----- 988
Db 118064 G C A G G T C A G C C C T A T A G C C A G A T C C A A G A G C C A T G C T C A A C T C T A C A G A T G T A C C 118123
```

```
QY 988 ----- 988
Db 118124 C C T C T C A A C T T A G T C A G A C T A C C C A C C C A C C C A C A G A G T G A T G T C C A C T C G G T T A 118183
QY 988 ----- 988
Db 118184 G C C A C G C C C C A T T G T T T G G T C C A T C C T T G A C C A G T T G C A A A G C C C G T C A C A G 118243
QY 988 ----- 988
Db 118244 G T C T C T T T T A G T C G T C C C T C C G C A G G C C C T G C T T T C A C G C T C T A T G G G T C G C 118303
QY 988 ----- 988
Db 118304 C C C T C T T A G C G T C T T C T A C T T C C G T C C A T C T T A C T C T G T T T C C G T C G T G G C T C C T 118363
QY 988 ----- 988
Db 118364 C C C C T T A T A G A C A T T T A G C C T A T C T G C C A G G C A C T T T T C A G C T G C T G T C A T C C A A T G 118423
QY 988 ----- 988
Db 118424 C T G A C C C A C T C A A A G C C C G T C C T T A A C A C T T A G A G T C C C T C T C A G C C T G G C C A C G C C 118483
QY 988 ----- 988
Db 118484 C A C A G G T C A C C A G A C A G C C T C T G A T C G C C T T C T G G G G C C G T G T A G T A G A A G T C T C C 118543
QY 988 ----- 988
Db 118544 T G T T C A C G C C C T T C T C A C C C A G A C T G G C C A T A C T T C A C A T T A G A C T C C C T C C C C 118603
QY 988 ----- 988
Db 118604 A G G C A C T G T T T C A C C C T G C C C A C T C C T C A C T G G C C A C C A C A C A C C T C T G C G C G C 118663
QY 989 -----Ty rAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTr 1006
Db 118664 T C T C A G G T A T G C C C G G A G T C C C T A T C T G A C A C A T C T T C C C G C C A A T C T A C G T G T G 118723
QY 1006 pSerPheGlyValValLeuTy rGluLeuPheThrTy rCysAspLysSerCysSerProSe 1026
Db 118724 G A G C T T C G A G T G T G T G T A C A G A C T C T T C A C C T A C T G C A C A A G A G C T G A G C C C A T C 118783
QY 1026 rAla- 1027
Db 118784 C G C - T G T G C G T C G C C T G C C C A T C C C T C G G T C C T C C C T T T G A T C T C C A A A T C C C C T C C T 118842
QY 1027 ----- 1027
Db 118843 G A C C T C T A G C C C C T A T C C T G A C C C C A G C C C T T C C T C C T G A C C T C C A G A T A T C T C C T G A C 118902
QY 1027 ----- 1027
Db 118903 C T C A G T C C T C C C T C G G A A T C C C A G C C C T C C T C C T C A T C T C C A G A T C C C T C T T G A C C C 118962
QY 1027 ----- 1027
Db 118963 C A G C C C T T C C C T G A C C C C C A G C C C T C C T C A T G A C G T C A G C C C C T T C C T G A C C C T A A C 119022
QY 1028 -----GluPheLeuArgMetMetMetGlyProGluAr 1037
Db 119023 A G T C T G C C C T A A C T T G C T C C T C C C A C A G A G T T C C T G A C A T G A T G G G G C C C T G A G C G 119082
QY 1037 gGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgLeuProPr 1057
Db 119083 T G A A G A C C C C C G C T C T C C G C C T C C T G A G C T G T G C A G A G G G C C C A G C C T C C C A C C 119142
QY 1057 oProProThrCysProThrGlu- 1064
Db 119143 A C C T C C C A C C T G C C C C A C C G A G G T G A G A G A G G A C T C A A G T T T C T C A G T T T C A A T T C T 119202
QY 1064 ----- 1064
```



```
Db      119203  TGATTGACAGACGCGTGCTGACTGGGAATACTTTTACAGGATGAGGCGCTCAGCTTC 119262
QY      1064  ----- 1064
Db      119263  AAGGGCTTCAGAAATGGTTTCACTTTTGTGAGCAAGGTCTTGCTATGTAGCCTTG 119322
QY      1064  ----- 1064
Db      119323  TCTGGCTGGAACGACATCCCTGCTGCTCAGTTTCTCTGACTGAGAGCAAAATGT 119382
QY      1064  ----- 1064
Db      119383  ATGTGCCACCAACACACCCCACTCCATATTTCATTTTCTTTTATTATTACTTAA 119442
QY      1064  ----- 1064
Db      119443  TTTATTTTACAGTCTGAGATTTGAAGCCAGGGTCTTACTCATGCTAGGCAAGCAGC 119502
QY      1064  ----- 1064
Db      119503  ACCCTGTGTCATGATAATTCGGGGATGGGGTTGGGGAAGGAGCGGCGCTTCCA 119562
QY      1064  ----- 1064
Db      119563  GGAGTGAATGGCGCCTATTGGAGAGCCAGGAGAGGTCAATCCAGTAGGAGACATGGCT 119622
QY      1064  ----- 1064
Db      119623  GTGGGCTGGGTGGAGTAAGAGACAGATGTCGCGGAGATGGGGGATGTGTGCTGTC 119682
QY      1064  ----- 1064
Db      119683  ACCCTAGCCATCTGACGGCCAGGCGAGGAGTGTGCCACGAGCAAACTAGTTCTGTTC 119742
QY      1064  ----- 1064
Db      119743  AGAGATCCATAAGGAGCCATGGAGAGCTGTGGAGAGAGTAGGGGCCATTAACTGACT 119802
QY      1065  -----ValGlnG1 1067
Db      119803  GGAACCCAGCCTCTCAGCTCAGCCACCCACCTTGCTTCCACCCAGCAGGTTTCAGGA 119862
QY      1067  uLeuMetGlnLeuCyATipAlaProGluProHisAspArgProAlaPheAlaThrLeuSe 1087
Db      119863  GCTCATGACCTGTGCTGGGCGCCAGCCGACGACCGCCAGCTTCGGACCCCTGAG 119922
QY      1087  rProGlnLeuAspProLeuTriArgGlyArgProGly 1099
Db      119923  CCCCCAGCTGGAGCGGCTGTGGCGTGGAGAGCCCGGA 119959

RESULT 10
AC073750
LOCUS
DEFINITION Mus musculus clone RP23-330D8, WORKING DRAFT SEQUENCE, 28 unordered
pieces.
AC073750
AC073750.1 GI:8810367
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 223734)
DOE Joint Genome Institute.
TITLE Sequencing of Mouse
JOURNAL Unpublished
2 (bases 1 to 223734)
DOE Joint Genome Institute.
AUTHORS Direct Submission
JOURNAL Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
```

```
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1863672
Center clone name: RPCI-23_330D8
-----
Summary Statistics
Consensus quality: 196566 bases at least Q40
Consensus quality: 208241 bases at least Q30
Consensus quality: 210965 bases at least Q20
Estimated insert size: 208000; agarose-fp estimation
Estimated insert size: 221034; sum-of-contigs estimation
Quality coverage: 8.01 in Q20 bases; agarose-fp estimation
Quality coverage: 7.54 in Q20 bases; sum-of-contigs estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 28 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
* 1 1027: contig of 1027 bp in length
* 1028 1127: gap of unknown length
* 1128 1261: contig of 1034 bp in length
* 2162 2261: gap of unknown length
* 2262 3709: contig of 1448 bp in length
* 3710 3809: gap of unknown length
* 3810 4864: contig of 1055 bp in length
* 4865 4964: gap of unknown length
* 4965 5985: contig of 1021 bp in length
* 5986 6086: gap of unknown length
* 6087 7791: contig of 1606 bp in length
* 7792 10317: contig of 2526 bp in length
* 10318 10417: gap of unknown length
* 10418 11756: contig of 1339 bp in length
* 11757 11856: gap of unknown length
* 11857 13223: contig of 1367 bp in length
* 13224 13324: gap of unknown length
* 13325 16114: contig of 2791 bp in length
* 16115 16214: gap of unknown length
* 16215 20778: contig of 4564 bp in length
* 20779 20878: gap of unknown length
* 20879 23553: contig of 2675 bp in length
* 23554 23654: gap of unknown length
* 23655 26667: contig of 3014 bp in length
* 26668 26767: gap of unknown length
* 26768 30222: contig of 3455 bp in length
* 30223 30322: gap of unknown length
* 30323 34999: contig of 4577 bp in length
* 34999 39009: gap of unknown length
* 39009 39709: contig of 4610 bp in length
* 39709 46418: gap of unknown length
* 46419 46518: gap of unknown length
* 46519 50029: contig of 3511 bp in length
* 50030 50129: gap of unknown length
* 50130 56982: contig of 6853 bp in length
* 56983 57082: gap of unknown length
* 57083 66535: contig of 9353 bp in length
* 66536 73174: gap of unknown length
* 73175 73274: contig of 6639 bp in length
* 73275 87407: gap of unknown length
* 87408 87507: contig of 14133 bp in length
* 87508 100238: contig of 12731 bp in length
* 100239 100338: gap of unknown length
* 100339 119196: contig of 18858 bp in length
* 119197 119296: gap of unknown length
* 119297 140119: contig of 20823 bp in length
```

\* 140120 140219: gap of unknown length  
 \* 140220 169227: contig of 29008 bp in length  
 \* 169228 189327: gap of unknown length  
 \* 169328 198979: contig of 29652 bp in length  
 \* 198980 199079: gap of unknown length  
 \* 199080 223734: contig of 24655 bp in length.

# FEATURES

Location/Qualifiers  
 1..223734  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /clone="RP23-330D8"  
 /clone\_lib="RPCI mouse BAC library 23"  
 BASE COUNT 57507 a 53195 c 54134 g 56203 t 2705 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,24e-239 Length: 223734  
 Score: 4091.50 Matches: 1072  
 Percent Similarity: 32.75% Conservative: 2  
 Best Local Similarity: 32.69% Mismatches: 24  
 Query Match: 69.82% Indels: 2188  
 DB: 2 Gaps: 24

US-09-397-967A-16 (1-1099) x AC073750 (1-223734)

QY	1	MetAlaProSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSer	20
Db	124160	ATGGCACTCCAAAGTAGGAGACACCTCTGATCCCTCAGGCGCTCTTGAGCCCTCATCC	124219
QY	21	SerGluAlaGlyAlaLeuHisValLeuProProArgGlyProGlyProProGlnArg	40
Db	124220	TCAGAGCAGGAGCCCTGCATGCTCTCTCTCCCGGGGACCTGGGCTCCCGAGCGA	124279
QY	41	LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla	60
Db	124280	TTGTCAATCTCTTTTGGGGACTACTTGGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC	124339
QY	61	CysglyIle-----	63
Db	124340	TGTGGTGA- GTGGGTCCCTAGATGACCATGCCATGTCATTAAGCAACATCTATTGAGAAAGG	124398
QY	63	-----	63
Db	124399	GGATGCTGGGTTTTTGAGGGGTGTGAATTTTGGTAAGGAGGATCATGGGATCAGCCCG	124458
QY	64	-----LeuProValTyrHis	68
Db	124459	TGTAGTTGACTTGGGTGAATCATGTGTCCTCCATCCCTTCAGGCATCCTGCCTGTTATCAT	124518
QY	69	SerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProProSerHisIlePhe	88
Db	124519	TCGCTTTTCGCTCTGGCCACTGAGGACTTCTCTGCTGGTTTCCCAAGCCACATCTTC	124578
QY	89	CysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg-----	103
Db	124579	TGCATAGAGGACGTGGACACTCAAGTCTTGTCTACAGCTACCGTAGGAAGTGAACCCC	124638
QY	103	-----	103
Db	124639	TTGATTTATCTTCAGGGTAGTATTCTATTTCAGTCTATAGTAGAGGCTCTTTGAAACCAGA	124698
QY	103	-----	103
Db	124699	GAGGTTGGGCTGTTTGGTGTAGGCCACACTGCATTCAGTACCCAACTTGGGCTCTGGG	124758
QY	104	-----PheTyrPheProAspTrp	109
Db	124759	AAGCCAGAGCTGAGTGGGCCACCCCATCTGAAACCCACAGCTTTTATTTCCCTGACTGG	124818
QY	110	PheGlyLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeu	129
Db	124819	TTTGGGCTGGACATGTACCCGCTTTGGGCTGGCGCAAGATTTTGACAGTGCCATCCTT	124878

QY	130	AspLeuHisValLeuGluHisLeuPheAla-----	139
Db	124879	GACTTACAGTTTTAGACATCTCTTTGCTCAGTTGGTATTCTTCGGTGGGAAGGCTC	124938
QY	139	-----	139
Db	124939	AACCCAGGATGGGGACCCCAAGGTAAACAAATCCAGCAGCCAGGACGAAACCTCACT	124998
QY	140	-----GlnHisArgSerAspLeuValSerGlyArgLeuProVa	152
Db	124999	AACACCCCTTCTGTGGCTGACCCAGCAGCACCGAGTACCTGTGTAGTGGGGCCCTCCCGGT	125058
QY	152	lGlyLeuSerMetLysGluGlnGlyPheLeuSerLeuAlaValLeuAspLeuAlaGl	172
Db	125059	GGGCTTAGCATGAAGGAGCAGGAGAGTTCTTGAGCTGGCGTGTGACATTGGCCCA	125118
QY	172	nMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrVal-----	188
Db	125119	GATGGCTCGTGAAGCAGGCCAGGCCAGGAGAGTCTCTGAAGACGCTCAGGTGAGGACC	125178
QY	188	-----	188
Db	125179	TCGGGCCACCCCTTCGTGGTCTCTTTGGCTGGGATCTAGGGCCCTATCGTTTCCCTT	125238
QY	188	-----	188
Db	125239	ACTGGAACACATACCTCTTCCCAACTGTGTGAGCCACCCTCTCCGTGAGCCTGCCCTG	125298
QY	188	-----	188
Db	125299	CCGAAAGGTCCCTATGCTGTCTGTGTGTCTTCGGGGCCCCCTACACCCCTTAAGAGC	125358
QY	189	-----SerTyrLysAlaCysLeuProProSerLeuArgAspValI	202
Db	125359	TCGGCGTGGCCCCCCCCACAGTTTACAAAGCTGTCTGCCGCCAGCCTGCGCATGTGA	125418
QY	202	leGlnGlyGlnAsnPheValThrArgArgArgIleArgThrValValLeuAlaLeuL	222
Db	125419	TCCAGGCCAGAACTTCGTGACACGAGCGCATCCGAGGACCGTGTCTTGGCGGTGC	125478
QY	222	euProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeuAsp	241
Db	125479	G-CGCTGTGTGCTCCAGCGCCGCTACCGCTCATGCGCAAGTATATTGAGAC	125537
QY	242	LeuGluArgLeuHisProAlaAlaThrGluThrPheArgValGlyLeuProGlyAla	261
Db	125538	CTGAGCGGCTACATCCAGCGGCCACCCAGAGACCTTCCGTGTGGGGCTCCCGGGCGCC	125597
QY	262	GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSerAsn	281
Db	125598	CAGGAGGAGCCGGGCTTCTGCGTGTGGCGGGGCAACAGCATCTCTCTGGAGCTCCGGG	125657
QY	282	AspGlu-----	283
Db	125658	GACCA-GGAGGTGTGGGGCTGGGCTTGAGAAAGCTGGTCCGCGGGGAGCGTGGCGGT	125716
QY	283	-----	283
Db	125717	TAGTAAAGAGCCCGTCGAACAATCAGGAAAGTGGTAGTCCGGGGCGGACTGATTGCA	125776
QY	283	-----	283
Db	125777	CTACGGGGCAGGGGGCGGGTCAATTAAAGGTGGTCCGGGGATGGAGCTACCCAGGGA	125836
QY	283	-----	283
Db	125837	CTGACCTGGGGCGGTGACTAGAAGGGTGTAGACTAGTAGGGGGCGGGACGACAGAG	125896
QY	283	-----	283
Db	125897	CGCATAGTATGGATTGGATCGGTGGGTGTGACCCGACCCAGGCTGTCTCAGTCCCTTAAC	125956

QY 283 ----- 283  
Db 125957 CTCGTCCCGGACCCCATCCACCCCTCCCTCCCGGTAAAGTGGAGGTGGAGGAG 126016  
QY 283 ----- 283  
Db 126017 AAGGGAGGCTGTGCTGGGTCAAGTCTGTAATCCTCTGGGTTCTGGCTGTCTCTTTCT 126076  
QY 283 ----- 283  
Db 126077 CTCCTTTTTTAATGGTTTTGGAGACAGGGTTTCTCTGTATAGTCTGGCTGTCTGTGA 126136  
QY 283 ----- 283  
Db 126137 ACTCACTTTGTAGACCAGGCTGGCCTGGAACCTCAGAAATCCGCCCTGCTCTCGA 126196  
QY 283 ----- 283  
Db 126197 GTGCTGTGATTAAGGCGTGGCCACTACACCAGCTACTGCTGTCTCTTCTCTCTCAG 126256  
QY 284 LeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaPro 303  
Db 126257 CTTTTCAGACCTTCTGTGACTTTCGGAAATCGTGTATGTACGATCAAGAGGCCCA 126316  
QY 304 ArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIle 323  
Db 126317 CGTGTGGGTCCGGCAGGGGAACACCGCTGGTCACCGTCACCAGGATGGACGGCCACATC 126376  
QY 324 Leu----- 324  
Db 126377 GTGTGGGTCTTGGCACCCCTTATTCTGTGTGTGATCTCTGGAAAGGGGTGCGTTCCGTT 126436  
QY 325 -----GluAlaGluPheProGlyLeuProGluAlaLeuSe 336  
Db 126437 GTTGTGAAGAGTTCTCTGGCAGGAGCGAGTTTCCGGGCTCCCTGAGGCGCTGTC 126496  
QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCys 356  
Db 126497 CTTCTGGCCCTTGTGATGGGTACTTCCGCTGATCTCGACTCCGAGCAATATTTCG 126556  
QY 356 sIysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIle 376  
Db 126557 CAAGGAGGTGGCGCGCGCGCTGCTGGAGGAGGGGAGCTGTGCATGGACCCAT 126616  
QY 376 ethr----- 377  
Db 126617 CACGTAGGAGCTAGTTGGGTCACAGTGGGATGGGCTGGGCAGCGAGGACTGTCTGG 126676  
QY 378 -----LeuAspPheAlaIleHisLysLeuLysAlaA 388  
Db 126677 GTTGGACCTGGCTGTCTTGGCCTAGGTTAGACTTTGCCATCCACAAGCTGAAGCCG 126736  
QY 388 lAglySerLeuProGlyThrTyrIleLeuArgSerProGlnAspTyrAspSerPheL 408  
Db 126737 CTGGCTCCCTCCAGGCACCTACATTCTCCGCGCAGCGCCGAGGACTATGACAGTTTC 126796  
QY 408 euLeuThrAlaCysVal----- 413  
Db 126797 TTCCTACGCGCTGCGT - CCAGGTTGGTTTACTGTGTGCCAGATGGCTTCTCGAGGA 126855  
QY 413 ----- 413  
Db 126856 AGTGACCTGAGCTGGGACAGAAAGAGCGTTTCTATTTTTCTTGAAGACACTGGCC 126915  
QY 413 ----- 413  
Db 126916 TGTAAATCTAGAAATTAAGAGACTGAGGACGATCAAGTTTACAGGCTGCTGAGCTAT 126975  
QY 413 ----- 413  
Db 126976 ATAGAGAAATATGTTTCTCATGGAGGAAAGGTTTTTTCACCAAGCTGAGGGTGT 127035  
QY 413 ----- 413

Db 127036 CTTTCATCCATGGCATTTCACAGCATGAGAGACAGAACTGACACCCATAAATCGCGTCTG 127095  
QY 413 ----- 413  
Db 127096 ACCTCCACAGTGCACGGTGACTTGTCATCTACACAGAAACACACATAAGTGAATGTAATTT 127155  
QY 413 ----- 413  
Db 127156 TAAGAACTCGAAAGAGCTCAGGAAACAGTCTCAGCATGCGGGGGGGGGGAGAGAAG 127215  
QY 413 ----- 413  
Db 127216 CATCCTTAGAGTTGGAGACAGAACTGAGCATACAGTCTTCCAGAAAGACGAGAAGGGGT 127275  
QY 414 -----GlnThrProLeuGlyP 419  
Db 127276 GGGTAGCTCTGCTCCACATCTGGCTCAGGTCCAGGTCTCTTTGACAGCTCCTCTTTGGCC 127335  
QY 419 rAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyL 439  
Db 127336 CGACTTACAAAGGCTGCCTCATCCGCCAGGACCCCAAGCGGGGCTTCTCCCTGGTTGGCC 127395  
QY 439 euSerGlnProHisArgSerLeuArgGluLeuAlaIaCysTyrAsnSerGlyLeuA 459  
Db 127396 TCAGCCAGCCCCACAGAAAGCTCGGGAGCTGCTTGCAGCCTGCTGGAATTTCTGGGCTGC 127455  
QY 459 rgValAspGlyAlaLeuTyrLeuThrSerCysAlaProArgProLys----- 476  
Db 127456 GAGTAGACGTCTGCTCCCTGAACTTAACTCCTGCTGCTCCAGACCCCAAGGTGAGT 127515  
QY 476 ----- 476  
Db 127516 CAGTTCCTCCCTAACCCAGCCCGCAGCTCTTGGGTGGTAGCAGAAACCAAGGACCATGG 127575  
QY 476 ----- 476  
Db 127576 CGATTAGGACAGAAACAGACTTTTGGCACAGTCTTCAGACATCAAGTCAGTTATTTT 127635  
QY 476 ----- 476  
Db 127636 CAGTTGGCGCGCTTTCTTCCCATGGGTGAATGCCACTAAGCGCAACCCCTGCTGGACA 127695  
QY 476 ----- 476  
Db 127696 CATCTCCCTCTGTGGTGTCTGGTTAAACAAAGACCAAGCGCGCTGGAGGAAACCCAGT 127755  
QY 476 ----- 476  
Db 127756 ACGTTCAATGTTCAAATCAGCCCCCAAGTCAGAAATCTCAGGAGCAGCTTCCCGGGCGGCC 127815  
QY 476 ----- 476  
Db 127816 CCAGGCTAAGTATGAGGAGTCGAAATCTTGGAAACCTTGGCTCTGACCTCTGCTTTCTT 127875  
QY 477 -----GluLysSerAsnLeuIleValArgArgGlyCysAsnProAlaProAlaP 494  
Db 127876 GCCATCCAGAAAGTCCAAATTTGATCGTGGTGGGAGGGGTGCACCCCGCGCTGCC 127935  
QY 494 rGlyCysSerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrA 514  
Db 127936 CTGGCTGTCTCCCGTCTCTGCTGTGGCTGTACACAGCTGAGCTTCCACAAATTCACACGG 127995  
QY 514 spSerLeuGluTrp----- 518  
Db 127996 ACAGCTGGAGTGGGTGAGGTTCCCGAGGTGGAAGCCAGGGCGGGCCCTTTTGGGAGG 128055  
QY 519 -----HisGluAsnLeu 522  
Db 128056 ACGTGAAGCCACCTTCAACTGTCTCGCTGGGTGTCTCCCGCCGACGACGAGAACCTG 128115  
QY 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgGluValValAspGlyGlu 542



QY 738 lnLeuProAlaLeuLysThrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrA 758  
Db 130334 AGCTGCCGCTCTCAAAATGACAGAACTGCGGACTTATCACAGTGCATGCGCTATG 130393  
QY 758 spProGlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThr 778  
Db 130394 ATCTGGCGCGGCCCTCTCTCCGAGCTATCTCTCAGAGACCTCAACGGCTCATATACAT 130453  
QY 778 ex----- 778  
Db 130454 CAGGTCCCGAGAGGGGGGGCGGAGGGCGGAGAGGGTCCCTGAGCTGGGTCCAGGAA 130513  
QY 778 ----- 778  
Db 130514 GTTTGCTGTTCTAGCAGTCAGCGTGAACCTAAGACAGCTGCTGCAACAACCTAAGACCC 130573  
QY 778 ----- 778  
Db 130574 GATCTGCATTTGTTTGCAGTGTAGCCAGGTTGGTGTATTTTACAATGTAGCCCAACTAG 130633  
QY 778 ----- 778  
Db 130634 CATCTCAGCTGGATTTCTCCCTGCTTACCATCATGGGTGCTGGGATTTGGTTTATTT 130693  
QY 778 ----- 778  
Db 130694 TGTATTCTTAAATTTTGTGCAAGATGTTCTCATCTGAGGATATCCAAAGTCACAGACT 130753  
QY 778 ----- 778  
Db 130754 GCGGTCACTGCATGGCTGTCTCTCGGGTCCCTTTTGTCTGATGCCGTTGGGGTCCCA 130813  
QY 778 ----- 778  
Db 130814 CAGAGTGTGCTTTGCCCGACAGTGGTGTCTCTCATTCGCCATTAAGTACTTTGGGGCG 130873  
QY 778 ----- 778  
Db 130874 TCTCATACGCACAGAGAGCAAGGTTCCAGGTGGAGTTTACACACCAATTTCAGTTTGGG 130933  
QY 779 -----AspTyrG 781  
Db 130934 ATTGAGGCAGAGACTTACCACTTACCCCACTTCTCACGTGATCCCGGCCCTAGATTACG 130993  
QY 781 luLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAlaG 801  
Db 130994 AGCTCTCTCAGACCCCAACACCTGGCATCCCGAGTCTCTCGAGATGAGCTGTGC--GGTG 131050  
QY 801 lyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrI 821  
Db 131051 GCGCCAGCTCTATGCTGCCAGGACCCCGCCATATTCGAGGAGAGACACCTTAAGTACA 131110  
QY 821 leSerLeuGlyLys----- 826  
Db 131111 TCTCTTTGCTGGCAAGTGTAGTGGCGGCGCATGTGGGGAGGAACTGGGTGGGTGGAT 131170  
QY 827 -----GlyAsnPheGlySerValGluLeu 834  
Db 131171 GGGTCAGGTGGACACTGCCCTCTCATCTCCACAGGGCACTTTGGCAGCGTGAGGCTG 131230  
QY 835 CysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln 854  
Db 131231 TGCCGCTATGACCCCTGGGGACAAATACGGGACCCCTGGTGGCAGTGAACAGACTACAG 131290  
QY 855 HisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeu 874  
Db 131291 CACAGCGGCCACAGACAGAGGAGACTTCCAGCGGAGATTTCAGATCCTTAAGGCTCTG 131350  
QY 875 HisSerAspPheIleValLysTyrArgGlyValSerTyrGlyPro----- 889  
Db 131351 CACAGCGACTTTCATCGTCAAGTACCAGGGGAGTCAAGTATGGGCCAGGTGAGCGGAGCAG 131410  
QY 889 ----- 889

Db 131411 CATCTCGGGNAACGGGTTCCAGTCCCCTTCTACCCCTTCCAGTAGGGCCCTCTTGAACA 131470  
QY 889 ----- 889  
Db 131471 AGTCCGTTAACTCCCATGAATCCAGCTTCTATAGCTGGAGGTTGAACGAATCAATACAC 131530  
QY 890 -----GlyArgGlnSerLeuArgLeu 896  
Db 131531 CCAAGTAGAGGTGCTCGAGAGCTACTGTCCGTGTTCCAGAGTCCGAGAGCCTGCGGTTG 131590  
QY 897 ValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly--- 915  
Db 131591 GTGATGGAGTACTGCTGCCAGCGGCTGCTTGCAGACTTCTCTGCAGCGCATTCGCGCGCGC 131650  
QY 916 LeuHisThrAspArgLeuLeuPheAlaTyrGlnIleCysLys----- 930  
Db 131651 CTGCACACCGACCGCCTACTGTCTGTTGCTTGGCAGATCTGCAAGGTGCGAGCATCGCG 131710  
QY 930 ----- 930  
Db 131711 AGATCGGATGGGTCTCTGGGGTCCGGATCCCGGGCCACGTGCTATCATGGCTCTCT 131770  
QY 931 ----GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAs 949  
Db 131771 GCAGGGCATGGAGTACTTGGTGGCGCGCTGCGGTACACCGTGACCTGGCTGGCGCGAA 131830  
QY 949 nIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeu 969  
Db 131831 CATCTGTGGAGAGCGAGGCTCATGTGAAGATCGCGACTTCGGCTTCGCTAAGCTGCT 131890  
QY 969 uProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTyr-- 988  
Db 131891 GCCCTGGAAAGGACTACTACGTGTCCGAGCTGCGCAAGCCCATCTTTTGGTA 131950  
QY 988 ----- 988  
Db 131951 CAAGACCAAGACTCCAGAGCTCACGCCCACTGAGGCCCACTCCCGCACTCTACAGCCTT 132010  
QY 988 ----- 988  
Db 132011 AAACCGTCCAGACCCAGCTCCACCTCAGTAGCTTTGCCCTCGCAAAGCCCTTGCCACA 132070  
QY 988 ----- 988  
Db 132071 GCAGGGTCAGCCCTATAGCCAGATCCAGAAAGCATGTCTCAACTCTACAGATGTACC 132130  
QY 988 ----- 988  
Db 132131 CCTCTCAACTTAGTCAGGACTACCCACCCACAGAGATGATGCTCCACTCGGTGTTA 132190  
QY 988 ----- 988  
Db 132191 GCCAGCCCCCATTTGTTCTGGCTCCATCTCTTGACCAAGTCTGAAAGCCCGTCACAG 132250  
QY 988 ----- 988  
Db 132251 GTCTCTTTTAGTCCCGTCCCTCCGAGGGCCCTGCCCTTTCACGCTCTATGGGGTCCGC 132310  
QY 988 ----- 988  
Db 132311 CCCTCTTAGCGTCTTCTACTTCTCGTGCATCTTACTCTGCTTTCCCGTCCGTGCTGCTCT 132370  
QY 988 ----- 988  
Db 132371 CCCCTTATGACATTTACGCTTATCTGCCAGGCACTTTTCCAGCTGCTGCTCATCCAATG 132430  
QY 988 ----- 988  
Db 132431 CTGACCCCACTAAAGCCCCGCTTAACTTAGAGTCCCTCTCTCAGCTGCGCCACGCC 132490  
QY 988 ----- 988

```
Db 132491 CACAGGTCACAGAGCCTCTGAGTCGCTTCTGGGGCGGTGCTAGTAGGAAGTCTCC 132550
QY 988 ----- 988
Db 132551 TGTTCACGGCCCTTCTCACCACAGAGTGGCCATATTCACATTGACGACTCCCTCCCC 132610
QY 988 ----- 988
Db 132611 AGCCCACTGTTTCACACCTGGCCACACTCTCCTCACTGGCCACACACACTCTGTCCGGC 132670
QY 989 -----TyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTr 1006
Db 132671 TCTCAGGTATGCCCGGAGTCCCTATCTGACAACTCTTCTCCCGCAATCTCAGCGTG 132730
QY 1006 pSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspIlySerCysSerProSe 1026
Db 132731 GAGCTTCGGAGTGTGTGTACGAGCTCTTCACTACTGCGACAAGAGCTGCAGCCCATC 132790
QY 1026 xAla----- 1027
Db 132791 CGC-TGTGCGTGGCTGCCCATCCCTCGGTCTCCCTCTTGATCTCCAAATCCCTCCT 132849
QY 1027 ----- 1027
Db 132850 GACCTCTAGCCCTATCTTGACCCAGCCCTTCTCCTGACCTCCAGATCATCTCCTGAC 132909
QY 1027 ----- 1027
Db 132910 CTCAGTCTCCTCCGGAATCCACGCCCTTCTCCTCATCTCCAGATCCCTTTTGACCC 132969
QY 1027 ----- 1027
Db 132970 CAGCCCTTCCCTGACCCCGAGCCCTTCTCATGAGTCAGCCCTTCTGACCCTAAC 133029
QY 1028 -----GluPheLeuArgMetMetGlyProGluAr 1037
Db 133030 AGTCTGCCCTAACTGTCTCCTCTCCACAGAGTCTCTGAGCATGATGGGCCCTGAGCG 133089
QY 1037 gGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgLeuProPr 1057
Db 133090 TGAAGGACCCCGCTCTGCGCCCTCTGAGCTCTGAGCTGTGCGAGGCGCGCCCTCCACC 133149
QY 1057 oProProThrCysProThrGlu----- 1064
Db 133150 ACCTCCACCTGCCCCACCGAGGTGAGGAGGAGCTCAAGTTCTCAGTTTCAATCT 133209
QY 1064 ----- 1064
Db 133210 TGATTGACGACGCGTGCCTGACTGGGATATTTACAGGATGAGGAGCGCTCAGCTTC 133269
QY 1064 ----- 1064
Db 133270 AAGGCTTCAGAAATGTGGTTTTCACCTTTTGTGAGGCAAGTCTTGTATGACCTTG 133329
QY 1064 ----- 1064
Db 133330 TCTGCCCTGGAACCTGCATCTCTGCTCAGTTCTCCTGACTGAGAGACAAATGT 133389
QY 1064 ----- 1064
Db 133390 ATGTGCCACCACACACCCACCCCACTCCATATTTCTTTTATTATTACTTA 133449
QY 1064 ----- 1064
Db 133450 TTTATTTTTTACGTGCTGAGATTTGAAGCCAGGCTTCTACTGCTAGGCAAGCAGCG 133509
QY 1064 ----- 1064
Db 133510 ACCCTGTCTATGATAATTGGGGGATGGGGTTCGGGAAGGAGCGCGCTTCCA 133569
QY 1064 ----- 1064
Db 133570 GGAGTGAATGGCGCTATTGGAGAGAGCCAGGAGAGGTCAATCCCACTAGGACATGGCT 133629
```

```
QY 1064 ----- 1064
Db 133630 GTGGCTGGGTGGAGTAAGAGAGCAGATGTGCGGAGATGGGGATGTGTGCTGTC 133689
QY 1064 ----- 1064
Db 133690 ACCCTAGCATCTGCAGGCCAGGCGCAGGAGTGTGCCACAGGCAAACTTAGTTTCTGTC 133749
QY 1064 ----- 1064
Db 133750 AGAGATCCATAAGAGGCCATGGAGAGCTGTGGAGAGAGTAGGGGCCATTAACTGAGT 133809
QY 1065 -----ValGlnG 1067
Db 133810 GGAACACACGCTCTCAGCTCAGGCACCCGCCACCTTGTCCACCACGAGGTTTCAGGA 133869
QY 1067 uLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSe 1087
Db 133870 GCTCATGCACTGTGCTGGCGCCAGCCGCGCAGCAGCCGACCTTCGGCACCCCTGAG 133929
QY 1087 rProGlnLeuAspProLeuTrpArgGlyArgProGly 1099
Db 133930 CCCCCAGCTGAGCGCTGTGGCTGGAGAGACCCCGA 133966
```

RESULT 11

MMU71201 10326 bp DNA linear ROD 15-JAN-1997

LOCUS Mus musculus protein tyrosine kinase JAK3 gene, complete cds.

DEFINITION U71201

ACCESSION U71201

VERSION U71201.1 GI:1778748

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 10326)

AUTHORS Kumar,A., Toscani,A., Rane,S. and Reddy,E.P.

TITLE Structural organization and chromosomal mapping of JAK3 locus

JOURNAL Oncogene 13 (9), 2009-2014 (1996)

MEDLINE 97088640

PUBMED 8934548

REFERENCE 2 (bases 1 to 10326)

AUTHORS Kumar,A. and Reddy,E.P.

TITLE Direct Submission

JOURNAL Submitted (18-SEP-1996) Fels Institute, Temple University, 3307 N. Broad St, Rm 339 AHB, Philadelphia, PA 19140, USA

FEATURES

source

1. .10326

/organism="Mus musculus"

/mol\_type="genomic DNA"

/strain="129/J"

/db\_xref="taxon:10090"

/chromosome="19"

/map="19p13.1-p13.2"

join(92..287,441..564,740..851,963..1108,1317..1599,2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,8524..8641,8907..9017,9707..10212)

/product="JAK3"

join(104..287,441..564,740..851,963..1108,1317..1599,2170..2289,2372..2529,2617..2728,3230..3413,3781..3902,3998..4129,4596..4680,5022..5149,5642..5774,5905..6056,6177..6327,6856..6995,7074..7260,7427..7551,7629..7801,8524..8641,8907..9017,9707..9772)

/function="protein tyrosine kinase"

/codon\_start=1

/product="JAK3"

/protein\_id="AAB40917.1"

/db\_xref="GI:1778749"

/translation="MAPSETPLIPORCSLSSEAGALHVLPPRGPPORLFSFGDYLAECLVRAAKACILPVYHSLFALATEDFSCWPFPSHFICIEDVDQVLVRL

RFYFDFWFLGTCRHFRLKDLTSAILDHLVLEHFAQHRSDLVSRGLPVLGSMKEQG  
EFLSLVDLDAQAREQAQRPGBELLKTVSKACLPPLSLRDIVOQONFVRRIRRTTV  
LALRVSPARPATATAGOVYLDLERLHPAATTFERVLGPAQEEPCGLLRVAGNGIS  
WSSGDQLFOTFCDFEIVDVSIKQPTCGSRHRLVTVTRMGHILEAFPGILPEAL  
SFVALVDGFRLLICDSRHYCKEVAAPRLLEEABUHGPFITLDFAIHLKAAASLP  
TYLRRSPQDYDFLUTACVOTPLGPDYKGLIRQDPSPAFSLVLSSTPEACGTCLQ  
CWNLSRVDGAALNLTSCCAPRKRKSNLIVRRGCTPAPGCSGCCALTSFHT  
IPTDSLGHENLGHGSFTKI PRGRREVVDGETHDSEVLLKVMDSRRNCMESFLEAA  
LMSQVYPHLVKGWCMAGDSIMVOEFVYLGALDMVLRKRHLVSAKWLQVTKOLA  
YALNLEDKGLPHGNYSARKVLLABSGDGNPPFIKLSDFGVSPVTLSEMLDRIPW  
VAPECLQEAQTLCLADKNGFGATTWEVFSGPAHITSLEPAKKUFIYEDQGLPALK  
WTELAGLITOCMAVDGRRPSFRALLDLNLGSLTSDYELSDPTGPPSPRDLCCGA  
QYACQADPAIFERHLKYLSLKGNGFSGVLCRYDPLDNTGPLVAVKLOHSGPQO  
RUFQRIQILKALHSDPIVKYRGVSPGSRULRVMEYLPSPGLDFLQRHAAUHT  
DLRLFLAWOIKQMEVILGARCVRDLAARNILVSEAHVKIADGLAKLLPLGKDYI  
VVREPGQSPIFWYAPSLSDNIERSQSDVNSFVYIYELPTYCDKSCSPSAEFLRMWG  
PRERGPLRLLELLABGRRLPPPTCTPEVQELMQLCVAPSRITQQPSAP"  
BASE COUNT 2088 a 3007 c 2880 g 2300 t 51 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1,15e-231 Length: 10326  
Score: 3935.00 Matches: 1059  
Percent Similarity: 32.79% Conservative: 6  
Best Local Similarity: 32.60% Mismatches: 31  
Query Match: 67.15% Indels: 2163  
DB: 10 Gaps: 25

US-09-397-967A-16 (1-1099) x MMU71201 (1-10326)

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
DB 104 ATGGCACTCCAAAGTGAGGAGACACCTTGATCCCTCAGCGCTCTTGACGCTCTCATCC 163  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
DB 164 TCAGAGCGAGGAGCCCTGATGTCCTCTCCCGGGGAGCTGGGCTCCCGAGCGA 223  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
DB 224 TTGTCAATCTCTTTGGGACTACTTTGGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC 283  
QY 61 Cys----- 61  
DB 284 TGTGTGAGTGGTCCCTAGATGACCATGCCATGTAAAGCAACATCTATGAGAAAGGG 343  
QY 61 ----- 61  
DB 344 ATGCTGGGGTTTGAGGGGTGTGGAATTTTGTGTAAGGAGGATCATGGGATCAGCCCGTG 403  
QY 62 -----GlyIleLeuProValTyrHisSer 59  
DB 404 TAGTTGACTTGGTGAAGTGTGCGCCATCCCTCAGGCATCTGCTCTTTATCATTCG 463  
QY 70 LeuPheAlaLeuAlaThrGluAspPheSerCysTyrPheProProSerHisIlePheCys 89  
DB 464 CTTTTCGCTCTGCGCACTGAGGACTCTCTGCTGCTTCCCGGAGCCACATCTTCTGC 523  
QY 90 IleGluAspValAspThrGlnValLeuValTyrArgLeuArg----- 103  
DB 524 ATAGAGGACGTGGACACTCAAGTCTTGTGCTACAGGCTACGCTAGGAGTGAACCCCTTG 583  
QY 103 ----- 103  
DB 584 ATTTATCTCAGGGTAGTATTTCTATTAGTCATAGTAGAGGCTCTTGAAACAGAGAGG 643  
QY 103 ----- 103  
DB 644 TTGGGCTGTTGCTGTAGGCCACACTGCACTTCACTGAGCCCAACTTGGCGTCTGGAAGC 703  
QY 104 -----PheTyrPheProAspTyrPheG1 111  
DB 704 CAGAGCTGAGTGGGCCACCCCACTTCTGAACCCACAGCTTTTATTTCCCTGACTGCTTGG 763

QY 111 YLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLe 131  
DB 764 GCTGGAGACATGTCACCGCTTTGGGTGCGCAAAAGATTTTGACAGTGCCTTCCTGACT 823  
QY 131 uHisValLeuGluHisLeuPheAla----- 139  
DB 824 ACATGTTTAAACAATCTCTTTGCTCAGGTGTGTATCTTCGGTNGGAAGGCTCAACCCA 883  
QY 139 ----- 139  
DB 884 GGGATGGGGGACCCCAAGGTAAACAATCCAGCACCCAGGACGGAACCTCACCTAACCA 943  
QY 140 -----GlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLe 154  
DB 944 CTTCTGTGGCCGACCCAGCACCGCAGTACCTGTGTGAGTGGGCGCTCCCGGTGGGCT 1003  
QY 154 uSerMetLysGluGlnGlyPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAl 174  
DB 1004 TAGCATGAAGGAGCAGGAGAGTTCCTGAGCCTGGCGCTGCTGGACTTGCGCCAGATGGC 1063  
QY 174 aArgGluGlnAlaGlnArgProGlyGluLeuLysThrVal----- 188  
DB 1064 TCGTGAGCAGGCCAGCGCCAGGAGCTGCTGAAGACGGTCAGGTGAGGGCCTCGGNA 1123  
QY 188 ----- 188  
DB 1124 CCGCTTCGTGTCTTCTTGGCTGCGGATCTAGGGCCCTATCGTTTTCCCTCCCGGAA 1183  
QY 188 ----- 188  
DB 1184 CACATACCTCTTCCCAACTTTGTGAGCCACCCTCTCCGTACAGCTGCCCGCGGAAAGGT 1243  
QY 188 ----- 188  
DB 1244 CCGCTATGCTGTCTCGTGTCTTCGGGCCCCCTCACCACCCTAAGAGCTCGGCNTGNN 1303  
QY 189 -----SerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyG 205  
DB 1304 CCCCCCCCCACAGTTACAAAGCCTGTCTCCCGCCAGCCTGGCGGATGTATCCAGGGCC 1363  
QY 205 InAsnPheValThrArgArgGilleArgThrValValLeuAlaLeuLeuProCysG 225  
DB 1364 AGAATCTGTGACACGAGCGCATCCGAGACCGGTGTGTGTGGCGCTGC--GCCTGTG 1421  
QY 225 lYArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgL 245  
DB 1422 GTCGCTGCCAGCGCCGACCGCTACGCTCATGSCCAAGTATAT-CTGGACCTGGAGCGGC 1480  
QY 245 euHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGluG 265  
DB 1481 TACATCCAGCGGCCACCCAGAGACTTCCGTGTGGGGTCCCGGGCGGCCAGGAGGAGC 1540  
QY 265 roGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrPheSerSerAsnAsp----- 282  
DB 1541 CGGGGCTTCTGCTGTGGCGGGGACAACGGCATCTCTCTGGAGCTCCGGGGGACCGAGAGG 1600  
QY 282 ----- 282  
DB 1601 TGCTGGGCTGGCTTGAGAGCGGGGTCCGCGGGGAGCGGTGGCGGTGTAGTAAAGAGC 1660  
QY 282 ----- 282  
DB 1661 CCGTCCGAACAATCAGAAAGTGGCTAGTCCCGGGCGGGCTGATTGCACCTACGGGGCAGG 1720  
QY 282 ----- 282  
DB 1721 GGGCGGGGTCAATTTTAAAGGTGGTCCGGGATGGAGCTGCCCCAGGAGCTGACCTGG 1780  
QY 282 ----- 282  
DB 1781 GCGTACTAGAGGTGAGACTAGTAGGGGGGGGAGCGGACAGACAGCGGATAGTAGTGG 1840











Db 8366 CTTCTGGGGGCGTGCTAGTAGGAAGTCTCTGTGTACGCCCCCTTCTCTCACCCCCAGAC 8425  
QY 990 ----- 990  
Db 8426 TGCCCATCTTACATTGACGACTCCCTCCCTGGCCACTGTTTTCACACCCCTGCCACTC 8485  
QY 991 -----ProGluSerLeuSer 995  
Db 8486 TCCTCACTGGCCACACACACCTCTGTCCGCTCTCAGGTATGCCCGGAGTCCCTATCT 8545  
QY 996 AspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeu 1015  
Db 8546 GACAAATCTTCTCCCGCCCAATCTGACGTGTGAGCTTCGGAGTGGTGTGTGACGAGCTC 8605  
QY 1016 PheThrTyrCysAspLysSerCysSerProSerAla----- 1027  
Db 8606 TTCACTACTGCGACAAAGAGCTGCAGCCCATCCGCTGTGCTGCGCTGCCCATCCCTCGG 8665  
QY 1027 ----- 1027  
Db 8666 TCCTCCCTCTTGATCTCCAAATCCCTNCTGACTCTCTAGNCCCTATCTCTGACCCCGCCC 8725  
QY 1027 ----- 1027  
Db 8726 TTCCTCCTGACCTCCAGATCATCTCCTGACCTCAGTCTCTCCCTCCGGAATCCCAAGCCCTT 8785  
QY 1027 ----- 1027  
Db 8786 CCTCCTCATCTCCAGATCCCTTCTTGACCCAGCCCTTCCCTGANNCCANCCCTTCTCT 8845  
QY 1027 ----- 1027  
Db 8846 CATGACNTCANCNCCCTTCTGACCCCTAACAGTCTNCCCTAACTTCTCTCTGCCACA 8905  
QY 1028 -GluPheLeuArgMetMetClyProGluArgGluGlyProProLeuCysArgLeuLeuG1 1047  
Db 8906 GGAGTTCCTGCGCATGATGGGGCTCAGCGTGAAGGACCCCTCTGCGGCTCTCTGGA 8965  
QY 1047 uLeuLeuAlaGluGlyArgArgLeuProProProThrCysProThrGluVal----- 1065  
Db 8966 GCTGTGGCAGAGGGCGCGAGCTCCACACCTCCACCTGCGCCACCGAGT-GAGGA 9024  
QY 1065 ----- 1065  
Db 9025 AGGAGACTCAAGTTCTCAGTTTCAATTCTTGATTGACAGCAGCGGCTCAGCTGGGA 9084  
QY 1065 ----- 1065  
Db 9085 ATACTTTACAGGATGAGGGAAGCTCAGCTTCAAGGCTTCAGAAATGTGGTTTTCACTTT 9144  
QY 1065 ----- 1065  
Db 9145 TGTNGAGGCAAGTCTTGCTATGTAGCTTGTGTGCTGGAACTCGACATCTCTCTGCC 9204  
QY 1065 ----- 1065  
Db 9205 TCAGTTTCTCTGAGCTGAGAAGCAAAATGTATGTGCCACCAACACCCACCCCACTCC 9264  
QY 1065 ----- 1065  
Db 9265 ATATTTCAAATTTCTTTTATTATTATTATTATTTTACAGTGTGAGATTGAAG 9324  
QY 1065 ----- 1065  
Db 9325 CCAGGCTCTACTCATGCTAGCAAGCAGCACCCTGTGTATGATAATTCGGGGATGG 9384  
QY 1065 ----- 1065  
Db 9385 GGGTTGGGAAGGAGGCGCGGTTCCAGAGTGAATGGCCCTATTGGAGGAGCCAGG 9444  
QY 1065 ----- 1065  
Db 9445 AGAGTCAATCCAGTAGGAGACATGGCTGTGGGCTGGGTGGAGTAAGAGAGCATGATG 9504

QY 1065 ----- 1065  
Db 9505 TGCNGCAGATGGGGATGTGTGCTGTCACTCCCTAGCCATCTGCAGGCCAGGAGGT 9564  
QY 1065 ----- 1065  
Db 9565 GTTGCCACAGGCAACTTAGTTTCTGTTCAGAGATCCATAAGGAGCCATGGAGAGCTGTGG 9624  
QY 1065 ----- 1065  
Db 9625 AGAAGAGTAGNNGNCATTAAACTGAATGGAACACCACTCTCAGCTCAGGCATCCGCCA 9684  
QY 1066 -----GlnGluLeuMetGlnLeuCysTrpAlaProGluPr 1077  
Db 9685 CCTTGATTCCTTCCACACACAGGTTCCAGGAGCTCATGCAGCTGTGGTGGCGCCAG-CC 9743  
QY 1077 oHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyAr 1097  
Db 9744 GCAGGACCGGCCAGCTTGGCACCTTGAGCCCCCAGCTGGAGCGCTGTGGCGTGAAG 9803  
QY 1097 gPro 1098  
Db 9804 ACCC 9807  
RESULT 12  
AC103459  
LOCUS  
DEFINITION  
AC103459  
AC103459.6 GI:30520536  
HTG: HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS  
Rattus norvegicus (Norway rat)  
SOURCE  
ORGANISM  
Rattus norvegicus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
REFERENCE  
1 (bases 1 to 216847)  
Muzny, D. Marie, Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Chen, Y.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,  
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
Gumaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,  
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, J.,  
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,  
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,  
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
Lorensuhewa, L., Louisleged, H., Lozano, R. J., Lu, X., Ma, J.,  
Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,  
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
Mawhiney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,  
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
Nwaekemele, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K.,  
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,  
Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Woodden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

# TITLE JOURNAL

REFERENCE  
Unpublished

2 (bases 1 to 216847)

Worley, K. C.

# TITLE JOURNAL

Direct Submission

Submitted (25-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 216847)

Rat Genome Sequencing Consortium.

# REFERENCE AUTHORS

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

# COMMENT

On May 10, 2003 this sequence version replaced gi:24953141. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHQO

Center clone name: CH230-30H1

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 199697 bases at least Q40

Consensus quality: 201698 bases at least Q30

Consensus quality: 203226 bases at least Q20

Estimated insert size: 208186; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

----- NOTE: This is a 'working draft' sequence. It currently

consists of 1 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

----- This sequence will be replaced

by the finished sequence as soon as it is available and

the accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

----- The accession number will be preserved.

/db\_xref="taxon:10116"

/clone="CH230-30H1"

85405..86754

/note="wgs contig"

208102..208763

/note="clone boundary

clone end:T7

site:EcORI

end sequence:BH291939"

215623..216847

/note="wgs end extension

clone end:T7"

BASE COUNT 47554 a 54739 c 54387 g 48109 t 12058 others

ORIGIN

Alignment Scores:

Pred. No.: 4e-215 Length: 216847

Score: 3692.00 Matches: 1012

Percent Similarity: 30.76% Conservatives: 24

Best Local Similarity: 30.05% Mismatches: 62

Query Match: 63.00% Indels: 2280

DB: 2 Gaps: 24

US-09-397-967A-16 (1-1099) x AC103459 (1-216847)

Qy 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20

Db 123069 ATGSCACCTCCAGCGAGGAGACACCTCTCATCTCTCAGCGCTCTTGTAGCCTCTCATCC 123128

Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProGlnArgGlyProGlnArg 40

Db 123129 TCAGAGCAGGAGCGCTTCATGTCTCTCTCCCGGGGACCTGGCGCTTCCCGAGCGA 123188

Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60

Db 123189 CTGTCATCTCTTTTGGGGACTCTTGGCTGAGACCTGTGTGCGGGCTGCCAAGGCC 123248

Qy 61 CysGlyIle----- 63

Db 123249 TGTGTGA-GTGGGTCCCTAGTAGACCGAGGCCACTTAGTAGGGCTGAGCAACATCTTGAG 123307

Qy 63 ----- 63

Db 123308 AAAAGGAGTGCTGGGGCTTTGCGGGATGCGTAGAAGTTTGTAAAGGGAGGATCATGGCT 123367

Qy 63 ----- 63

Db 123368 CAGCTGTGTTGTCATTGGTGTAAGTGGGTCATCTCTTTCAGGCATC 123427

Qy 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTyrPhePro 83

Db 123428 CTGCGCGTTTATCATTCGCTCTCGCTCTGCGCACCGAGGACTTGTCTTGTGTGTTCCCC 123487

Qy 84 ProSerHisIlePheCysIleGluAspValPheThrGlnValLeuValTyrArgLeuArg 103

Db 123488 CCAAGCCACATCTTTCATAGAGAGCGTGGACACTCAAGTCTTGGTCTACAGGCTACGG 123547

Qy 103 ----- 103

Db 123548 TAGGAAGTGAGCCCTCTATTACCTTCAGGGTAGTATTCTTCTCTTTTCAAAGAT 123607

Qy 103 ----- 103

Db 123608 TTGTTTATTATTATATATAAATCAATTGTTGCTGTCTTCAGACACACGAGAGAGGCA 123667

Qy 103 ----- 103

Db 123668 TCAGATCTCATTTATAGATGGTTAAGCCACCATGTGTTATTTGGATTTCAGACTCAGGACC 123727

Qy 103 ----- 103

Db 123728 TCTGGAAGAGCAATCAGTGCTCTCACTGCCCATCTCTCCAGCCCCCAGAGTAGTATTCTT 123787

# FEATURES source

1..216847

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

```
QY 103 ----- 103
Db 123788 ATTCACTTAAGTAGAGGATATTGAGACCAGAGAGGTGGGCTGTTGCTATAGTAGGC 123847
QY 103 ----- 103
Db 123848 CACACAGCACTTCAGTGAACCAACTTGGACTCTGGGAAGTCAACAGCTGAGTGGACACCC 123907
QY 104 ----- PheTyrPheProAspTyrPheGlyLeuGluThrCysHisArg 117
Db 123908 CATTCTGGAACCCACAGCTTTTATTTCTGGTGGTTTGGACTGGAGACCTGTACCCG 123967
QY 118 PheGlyLeuArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeu 137
Db 123968 TTTGGGTTACACAAAGATTGACCAAGTGCATCTTGGACGTACATGTTTGAACACCTC 124027
QY 138 PheAla----- 139
Db 124028 TTTGC-CCAGGTTGGTATTCTTCGGTGGGAAGGCTCAACTAGGATGGGGACCCCAA 124086
QY 140 -----Gln 140
Db 124087 GSCAACATCTCAGCTCCAGGACAGAACCCACCTAACCAAAATCTTGTGGCAACCCAG 124146
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIysGluGlnGly 160
Db 124147 CACCCAGTGCCTGGTGAAGTGGGCTCTCCAGTAGGCTTTAGCCTGAAGGACCCAGGA 124206
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 124207 GAGTTCTCGAGCCTAGCTGTGTGACCTGGCCAGATGGCTCGTAAGCAGGCCCCAGCG 124266
QY 181 ProGlyGluLeuLeuIysThrVal----- 188
Db 124267 CCAGGAGAGCTGCTGAAGTCCGTCAAGTCAGGCGGAGGCGGCCAACCTCCCGTGGTCTCT 124326
QY 188 ----- 188
Db 124327 TTGGCTGGGACCTGGGGCCCTATCGTTTCCCTTACTGAACCTCTACAAGTACCTCCTT 124386
QY 188 ----- 188
Db 124387 CCCAACGTCAACTCTGCAGCCACCCTGCCAAAGGTCCCGATGTATGTCATGTGTC 124446
QY 189 -----SerTyrIysAla 192
Db 124447 CTTCCGGTCCCCTTACACCCCTAAGAGCTCGCGCTGCCGCCACCACCAAGTTACAAAGCC 124506
QY 193 CysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArg 212
Db 124507 TGTCTGCCGCCAGCTCGCGACCTGATCCAGGGCCAGAGCTTCGTGACGCGCAGGCGC 124566
QY 213 IleArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgPro-Ty 232
Db 124567 ATCCGACAGCCGTGGTCCAGCGCGTGG-CCGTGTAGTAGCTGCCAAGCCGACCGCTA 124625
QY 232 rAlaLeuMetAlaIysTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrGln 252
Db 124626 CGCACTCATGCCAAAGTACATCTTAGACTTAGCGGTGATCCAGGGCCACCACTGA 124685
QY 252 uThrPheArgValGlyLeuProGlyAlaGlnGluProGlyLeuLeuArgValAlaGln 272
Db 124686 GTCCTTCTCGTGGGACTCCCGGGCCGACAGGAGCGGGTTGCCCTGCGTGTGACAGG 124745
QY 272 yAspAsnGlyIleProTyrSerSerAsnAspGlu----- 283
Db 124746 GGATAATGGCATCGCTGGAGCTCCAGGACCA-GGAGGTTCTGGGGCAGGACTTGAGG 124804
QY 283 ----- 283
Db 124805 AGCGGGTCTCGGGGAGGCGTGGCAATTAGTAAGAGGCAATCGGGAGGAACCAATCA 124864
QY 283 ----- 283
Db 124865 GGAAGGGGCCAGTCGGGCGAGGCTGGCTGCACCTACCGGGTGGGTTAAATTAGAGCGG 124924
QY 283 ----- 283
Db 124925 ATCCGGGGTGGAGCTTACCCAGGACTGGACCTGGGGCGGTGACCAAGGTTGGGACTAG 124984
QY 283 ----- 283
Db 124985 ATAGCGCGGGCCGGCGCGGACAGGGATAGTCAGGATTGGTGGTGGCGGTGACTTGAC 125044
QY 283 ----- 283
Db 125045 GAGGCCAGCTTCATTAGTCCCTTTCCCTGCCCGGGAAGCTGTGTGGATCCAGGACTGT 125104
QY 284 -----LeuPheGlnThrPheCysAs 290
Db 125105 AATCCCCGACCTTAGCTTGTCTGTCTTCTTTCAGCTTTTCCACACCTTCTGTGA 125164
QY 290 PheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGln 310
Db 125165 CTTTCCAGAAATCGTGGATGTGAGCATCAAGCAGGCTCCACGCGTGGTCCGCGCAGGGA 125224
QY 310 uHisArgLeuValThrValThrArgMetAspGlyHisIleLeu----- 324
Db 125225 GCACCGGCTAGTCACCATCACCGAGATGAGCATCTCTGGTGGGTCTTGGCACCCG 125284
QY 324 ----- 324
Db 125285 TTTTCTGTGTGCTGACCTCGGAGCGGGTGGCTCCATTGTTACTGAAGATTCTCTCTG 125344
QY 325 -----GluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspG 343
Db 125345 CGCAGAAACGGAGTTCGCGGGCTGCGTGGCGCTGTCTTTCGTGGCGCTCGTGGATG 125404
QY 343 lYTrPheArgLeuIleCysAspSerArgHisTyrPheCysIysGluValAlaProProA 363
Db 125405 GGTACTTTTCGCTGATCTGCGACTCCAGGCATTTCTTCTGCAAGAGGTGGCACGCCAC 125464
QY 363 rgLeuLeuGluGluAlaAspValCysHisGlyProIleThr----- 377
Db 125465 GGTCTGTGAGAAAGAGCGGAGCTGTGCCATGGACCCATCACGTAAGGAGCTGCCCGG 125524
QY 377 ----- 377
Db 125525 GTCACAGAGGGAACGGGTTGGCGACGGGAACCTGTCTGGTGTGAACCTGGCTGTCTCC 125584
QY 378 -----LeuAspPheAlaIleHisIysLeuLysAlaAlaGlySerLeuProGlyThr 394
Db 125585 TGGCCAGGTTAGACTTTGGCATCCCAAGCTGAAGGCTGTGGCTCCCTCCCGGGCTCC 125644
QY 395 TyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysVal--- 413
Db 125645 TACATTCTCCGACGACGCCCGCAGGACTATGACAGCTTCTTCTTACTTCCTGTGTCCAG 125704
QY 413 ----- 413
Db 125705 GTTGGTTCACTGCTGTGTCAGGATGGCTTCTCTGAAGAGGTGACACCTGAATGGGACA 125764
QY 413 ----- 413
Db 125765 GAAAGGATGTTTCCATTTTTCCTTGAAGAGCACTGGCCTATAATCCTAGAAATTAAGAG 125824
QY 413 ----- 413
Db 125825 ACTGAGCAGGATCAAGCTTACGGCCAGGCTGAACATACATAGAGAAATATGACCCAAA 125884
QY 413 ----- 413
Db 125885 AATTGTGTGTGTGTGTGGGGGGGAGTGTCTCATGGGGGAAAGGTTTTTTCACCACCA 125944
QY 413 ----- 413
```



```
QY 617 SerTrpLysLeuGlnValThrLysGlnLeuAlaIleValLeuAsnTyrLeu----- 633
Db 128164 AGCTGAAACTGCAGGTGACCAAGCAGCTGGCCCTATGCCCTTAACCTACTTT-GGTGAGTGT 128222
QY 633 ----- 633
Db 128223 TCCTGTGTCCGTGTTTATTTCCTCTACCAATGAGGAAATGAAGAGTAGGCTCTGTAGTTT 128282
QY 633 ----- 633
Db 128283 GTATAGGAGCTCGGTGAGGACTGAAGCAAGAAACAGAGATGTGAGCCTGTTACGGGGTTG 128342
QY 633 ----- 633
Db 128343 AGAGCATCAAGGACTGTAAACCCTATGGGTGTCGAGGCCAGGCTGTGGGGACAGGGCATA 128402
QY 633 ----- 633
Db 128403 GAGATGGAAACTGGTAAGAGGAGAGGGTGTCTTGAATGGTGAATTTAGAGGCTTTAACTG 128462
QY 633 ----- 633
Db 128463 TCCCTAAAGCCATGGAAAGGTAGAGAAAGGGGGCTGGTGAAGGGTAGGTATGATTAG 128522
QY 633 ----- 633
Db 128523 TGTAAGGGCAGGTTAGGGTCACAAGACATGTCTTTCATGGACTTACAGTGACTGTCAAAGC 128582
QY 633 ----- 633
Db 128583 CTGTTCTGTCTGGCAGCATCAGGTAGGCTTTGAAGGATGCAGAGAGTTTGCTTAATAG 128642
QY 634 -----GluAspLys 636
Db 128643 TGTTTCATCAAACTCATCCACACTTCCTGTCTCTGAAACACTTAGAGGAGCAAA 128702
QY 637 GlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyAsp 656
Db 128703 GGCCCTCCCTCATGGCAACGCTCAGCGAGGAAGGTGCTGCTAGCTCGTCAGGGGGTGTAT 128762
QY 657 GlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676
Db 128763 GGAATCCACCTTTTCATCAAGCTGATGATCCTGGTGTGAGCCCACTGTTTGTAGCCTG 128822
QY 677 Glu----- 677
Db 128823 GAAAGTGTACAGCCAGCCTCACCCAACTATACCTGTGTCTGTCTGTGCACACACA 128882
QY 677 ----- 677
Db 128883 GTTACACAGGGCCTGCAGTGGCAAGAGACCCGCTGTGTCAGAGGTGACCGTACCATT 128942
QY 678 -----MetLeuThrAspArgIleProThrValAlaProGluCysLe 691
Db 128943 CTCACCTCCACATCTCCAGTGCTCACTGACAGGATTCCTGGGTGGCCCGGAGTGCT 129002
QY 691 uGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTr 711
Db 129003 CCAGGAAGCTGGAACGCTCACTTGAGGCTGACAAGTGGGGCTTTGGCGCCACCACCTG 129062
QY 711 pGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLys----- 728
Db 129063 GGAGGTGTTACGGGGGACCCCATGCACATTACCTCGCTGGAGCCCGCCAA--GGTCAGAG 129121
QY 728 ----- 728
Db 129122 CAGGTCTCAGCCTTATGATCTGGTGTGGTGGGGTAGGGTTAAACCACGTGACGGCAT 129181
QY 729 -----LysLeuLysPheTyrGluAspGlnGly 738
Db 129182 CTCAAACCCAAACGAAACACTGCCCTTCAGAAAGCTGAAGTTCTATGAGGACCGGGGACA 129241
QY 738 nLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAs 758
```

```
Db 129242 GCTGCGCGCCCTCAATATGACAGAGCTGGAGGGGCTCATCGACAGTGCATGCGTATGA 129301
QY 758 pProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSe 778
Db 129302 TCCTGCGCGCGCCCTCTCTCCGAGGCCATCCTCAGAGACCTCAACGGCCTCTATTACATC 129361
QY 778 ----- 778
Db 129362 -AGGTACCTCGGGGTGGAGAGGGTCCCGAGCTGGGTCCCGAAGGTGTGTGCTCCAGC 129420
QY 778 ----- 778
Db 129421 AGTCAATGTGAGCTAAGACAGCAGCTGCTGGAACCTAACCTGTGGCTAAGATAAACTCAATCA 129480
QY 778 ----- 778
Db 129481 GCATTTGTTTGCATATGTAGTCCAGGCTAGTGTGTTTACAGTGTAGCCAGGTTGGCCTCT 129540
QY 778 ----- 778
Db 129541 CAGACTGGATTTCTCTCTCCCTCATCTGCTCTGGAATTTGTTTATTTGTTATCCGGGTC 129600
QY 778 ----- 778
Db 129601 ACAGGCTGGTGTCACTGTATGGCTCTGCTCTCGGGTCCCTTTGTTGTCTGTAGCCCTTG 129660
QY 778 ----- 778
Db 129661 GGATCCACAGAGTGTGCTTTTCCCGGACTGTGATCGTTCCACTCCCAACAAAGTACC 129720
QY 778 ----- 778
Db 129721 CGTGGCATCTCATATAATGCACAGAGAGACAAAGTTTCAGATGGAGTTCACACCAATTC 129780
QY 779 -----AspTyr 780
Db 129781 AGTTTGGGGTTTCAGGCACAGAAAGATACCCCACTACTCATGTGACCCCGCCCTAGATT 129840
QY 780 rGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAl 800
Db 129841 CGAGCTCTCTCAGACCCCAACACCTGGCATCCGAATCCCGAGATGAGCTGTGT--GG 129897
QY 800 aGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTy 820
Db 129898 TGGTCCCGAGCTCTATGCTTGCAGACCCCGCCCATCTTCGAGGAGAGACACCTTAAAGTA 129957
QY 820 rIleSerLeuLeuGlyLys----- 826
Db 129958 CATCTCCCTGTGGGCAA--GGTGTGTGAGCGGAGCTTGTGGAGCGAGGCTTAAGTTGGCG 130016
QY 827 -----GlyAsnPheGlySerVa 832
Db 130017 GGTGGAGGGTCAGGCAGAAACTGCGCTCTCTTCTCTCCACAGAGGGGAACTTCGCGCAGCT 130076
QY 832 lGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGl 852
Db 130077 GGAGCTGTGCGCTACAGCCCTCTGGGAGACAATAACGGGACCCCTTGTAGTAGTGAAGCA 130136
QY 852 nLeuGlnHisSerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLys 872
Db 130137 GTTACAGCAGCGGGCCAGAACAGCAGAGGAGCTTCAGCGAGAGATTTCAGATCTCTCAA 130196
QY 872 sAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyPro----- 889
Db 130197 GGCTCTACACTGTGACTTTCATCGTCAAGTACCGGGAGTCACTATGGGCCCTGTGTGAGAG 130256
QY 889 ----- 889
Db 130257 GCGCAGCAGCTCGGGAAGGGGTTTGAGTCCCGCTTCTACCCCTTTCCAGTTGGGGAGTG 130316
QY 889 ----- 889
```



```
Db 130317 TTGAGCAAGGCGTTAAACCCCATGAATCTCAGCTTCCATAGCTGGGGCTTGAGCGAAC 130376
QY 890 -----GlyArgGlnSerLe 894
Db 130377 AATGCACCGGGTAAAGGTGCTCGAGAGCTACTTTCCCTATCCGCGAGGTGCGCAGAGCCT 130436
QY 894 uArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisAr 914
Db 130437 GCGGTTAGTGATGGAGTACCTGCCAGCGCTGCCTGCGGAGCTTCCTGAGCGCCACCG 130496
QY 914 gGly----LeuHisThrAspArgLeuLeuPheAlaTyrGlnIleCysLys----- 930
Db 130497 CCGCGCGCTGCACAAACGACCGCTACTGCTGTTGCGCTGCGCAGATCTGCAAGGTGCGAGC 130556
QY 930 ----- 930
Db 130557 AGCGTGAGATCCGGATGGGGAATCTGGGGTCCGTGATCCCGGCCCCACATGCTATCACG 130616
QY 931 -----GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaA 947
Db 130617 GCCCCCTACAGGGCATGGAGTACCTGGCGCACGCCCTGCGTGATCGTGGCTG 130676
QY 947 laArgAsnIleLeuValIleuSerGluAlaHisValIleAlaAspPheGlyLeuAlaL 967
Db 130677 CCGCAACATCCTGTTGGAGAGTGAAGCATGTGAAGATTGCTGACTTCGGGCTCGCTA 130736
QY 967 ysLeuLeuProLeuGlyIleAspTyrTyrValValArgGluProGlyGlnSerProIleP 987
Db 130737 AGCTGTGCCTCTCGGAAAGACTACTACGTGTGTCCGCGAGCGCGGCCAAAGTCCCATCT 130796
QY 987 heTrp----- 988
Db 130797 TCTGGTATAAGACCAAAACCCCTGGGTTGGGGAATTTAGCTCAGTGGTAGAGCGTTGCCT 130856
QY 988 ----- 988
Db 130857 AGGAAGCGCAAGGCCCTGGGTTGCGTCCCGAGCTCCGAAAAAGAAAAAGAAAAA 130916
QY 988 ----- 988
Db 130917 GACCAAAACCCCTGACGCCACGCCCACTGAGCCCCAGACTCCACACTTTACTGCCTCAA 130976
QY 988 ----- 988
Db 130977 ATCGGTCCACTTCTCAAGCTTGCGCTCGCAGAGCGCGCCACAGCAGGGTCAGCCCC 131036
QY 988 ----- 988
Db 131037 GTAGCCAGCCACAGAGCCACGCCACAGCTTTATACAGTACCAGCCTAACTGTGTA 131096
QY 988 ----- 988
Db 131097 GTCGGGACTACCGACCCCTTCCACAGAGTGATGTTTAACTCATTTCTGGCTTCGC 131156
QY 988 ----- 988
Db 131157 CCTCAGTGACCACTGACAGCAAGCCCTTCACAGTCTCTTCTAGCTCGGCCCCCTCCGCA 131216
QY 988 ----- 988
Db 131217 CGGCGCGGCTTTACAGCTCCATGGGGTTGCGGCCCTCTAAGTCAACCCACTTCTGTCTAT 131276
QY 988 ----- 988
Db 131277 CTTTCTCGCTTTCCGCTGCTGGTTCCTCTCTTATGGAACCTTATACGTACCTGCATGC 131336
QY 988 ----- 988
Db 131337 CTATCTGCCCCACGACCTTTCCAGTGTGACATCATCCAAAGCCCGTCCCAATCCACTTA 131396
QY 988 ----- 988
Db 131397 GAGTCTCACCCCTCAACCTGGCCACGCCACAGGTCCCGCAGACAGCTCTGTGAGCGCCCTTA 131456
```

---

```
QY 988 ----- 988
Db 131457 TGCAGTCTTGGCTTCTAGTCTCTGTGACACACCCCTTCTCTAGTCCCCCAACCTAAC 131516
QY 988 ----- 988
Db 131517 ACACCCGTGTCCAGATCTTACTGTGCGCACACCCCTTTTCTGACTACGCTCTGTCGCGCTC 131576
QY 989 -----TyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyr 1007
Db 131577 TCAGGTACGCGCCCTGAGTCCCTATCTGACAAACATCTTCTCCGCCAATCTGAGCTGTGGA 131636
QY 1007 erPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerA 1027
Db 131637 GCTTCGGCTGGTGTGTATGAGCTCTTCCACTACAGCGCAAGAGCTGCGAGCCCATCCA 131696
QY 1027 la----- 1027
Db 131697 C-TGTGCGTGCAGGTCCCATCTCTCCCTCAGTCTTCTCTGACCTTAACCCCTCTCTT 131755
QY 1027 ----- 1027
Db 131756 GACCTCCAAACTTTCTCTTGTCTTCAGCGCTTCTCTCTGACCTCAGCCCTTCTTCCCA 131815
QY 1027 ----- 1027
Db 131816 ACCTCCAGTCTTCTCTCTGACCTCCAGACCACTTCTTGTGACCTCAACCCCTTCTCTCTGA 131875
QY 1027 ----- 1027
Db 131876 CCTCCAGTCCCTTCTCTCTGACTTCCAGCGCCCTGTCTCTGACCCAGCCCTCTCTTCTCTGA 131935
QY 1027 ----- 1027
Db 131936 GTCACAGCTCTTTTTTTTAAAGATTATTCATTTATTATATATATAGTACTGTAGCT 131995
QY 1027 ----- 1027
Db 131996 GTCCTCAGATACACCAAGAGGGCATCGGATCTTTTACAGATGATTTGTGAGCCACCAT 132055
QY 1027 ----- 1027
Db 132056 GTGGTGTGTGGAATTGAACCTCAGGACCTCTGGAAGAGAGCTCGGGGCTCTTAACCGATG 132115
QY 1027 ----- 1027
Db 132116 GGCCATCTTCCAGCCCGAGTCCAGCTCTTTTCTCTCACTCCAGATCCCTTCTCTGAC 132175
QY 1027 ----- 1027
Db 132176 CCCCACCTGTCTCTGACCCCTTCAGGCTTCTCTTATGTTATGATCAGCCCTTCTCTGACCC 132235
QY 1028 -----GluPheLeuArgMetM 1033
Db 132236 CAGCCCTCCCTCTGTCTGCTGCCCTTAACCTTGTCTCTCTCCACAGAGTCTCTGCGCATGA 132295
QY 1033 etGlyProGluArgGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyA 1053
Db 132296 TGGGGCTGAGCTGAGGGATCCCGGCTCTGCCACCTCTTGGAGCTGTGGCAGAGGGCC 132355
QY 1053 rgArgLeuProProProProThrCysProThrGlu----- 1064
Db 132356 GCGGTCTCCACCCACCTCCCTCCCTGCTCTACTGAGGTGAGGGAGACAGACTCAAGTTTCC 132415
QY 1064 ----- 1064
Db 132416 CAGTTTCCATCTCTGATTGACAGCATGGTTCATGACTGGGAGGCTTTTACATGAGGAGC 132475
QY 1064 ----- 1064
Db 132476 TCAGCTTCAAGGGCTTCAGAAATGNGGTTTTTTCACTTGTGTGAGGCAAAAGTCTTGTATG 132535
```







```

Db 2635 GAGCTGTGGCCGTGAGAGCTGCACAGGATTCGCCAAGGAGCTGCAGGACTTTGAG 2694
QY 866 ArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyVal 885
Db 2695 AGGAGATCCAGATCTGCTGCTGCACACGACTTCATGTCAGATTCAGGCGGCTC 2754
QY 886 SerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCys 905
Db 2755 TGCTACAGCGCTGGCGCGCGGCTGCGGTGGTGATGAGTATCTGCTGACGCGCTGC 2814
QY 906 LeuArgAspLeuLeuGlnArg-----HisArgGlyLeuHisThrAspArgLeuLeuLeu 923
Db 2815 CTGAGGAGTACCTGCAAAAGACACAGCACCGC---CTGAGACACCGCAGCTGCTGCTC 2871
QY 924 PheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArg 943
Db 2872 TAGCGCTGCAGATCTGCAAGGATGAGTACTCTGGGGCGCAGCGCTCGTGGCACCGC 2931
QY 944 AspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPhe 963
Db 2932 GACTTGGCCAGGAGAACATCTGTGGAGAGCGAGACCCACGTCAGATCGGTGACTTC 2991
QY 964 GlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGln 983
Db 2992 GGCTGGCCCAAGCTGTGCCACAGGACAAAGGACTACTACGTGTGTCAGGAGCCCGGCGAG 3051
QY 984 SerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSer 1003
Db 3052 AGCCCGCTGTTCTGTACGACCCGAGTCTGTGCTGACACGCTCTTCCCGAGCATCC 3111
QY 1004 AspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCys 1023
Db 3112 GACGTGTGAGCTGTGGGGTGTGCTCTATGAGCTCTTCCACCTACAGCAATAAGACGAG 3171
QY 1024 SerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCys 1043
Db 3172 AGCCCTCGGAGAGTCTTCCACATGATGGGCGCTCGAGAAACACACAGATCATCTGC 3231
QY 1044 ArgLeuLeuGluLeuAlaGluArgArgGluLeuProProProProProThr 1063
Db 3232 CATTTGTGGAGCTGTGAGGAGCTCCCGGGCTCCCGGCTCCCGGGCTGCCCATG 3291
QY 1064 GluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPhe 1083
Db 3292 GAGGTGTATGCGATGATGCTGAGCTGTGCGGTTCGCCCGCCAGTGCCAGACCCCTTC 3351
QY 1084 AlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArg 1097
Db 3352 ACCGAGTTGGCAGCCAGGTTGAGGCACTGCGGGATGGCGGT 3393

RESULT 14
AX203084 2914 bp DNA linear PAT 30-AUG-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0152892.
ACCESSION AX203084
VERSION AX203084.1 GI:15392441
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Vasilos,G.
TITLE Jak/stat pathway inhibitors and the uses thereof
JOURNAL Patent: WO 0152892-A 5 26-JUL-2001;
GENZYME CORPORATION (US)
FEATURES
Location/Qualifiers
1..2914
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 566 a 907 c 809 g 632 t
ORIGIN

```

## Alignment Scores:

Pred. No.: 5,8e-189 Length: 2914  
 Score: 3228.50 Matches: 614  
 Percent Similarity: 87.63% Conservatives: 45  
 Best Local Similarity: 81.65% Mismatches: 90  
 Query Match: 55.09% Indels: 3  
 DB: 6 Gaps: 3

US-09-397-967A-16 (1-1099) x AX203084 (1-2914)

```

QY 350 AspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGluAla 369
Db 5 GACTCCAGCACTTCTTCGAAAGAGGTGGACCCGCGAGGCTGTGGAGAGTGGCC 64
QY 370 AspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGly 389
Db 65 GAGCAGTGCACGCGCCCATCACTCTGGACTTTGCCATCAACAAGCTCAAGACTGGGGC 124
QY 390 SerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeu 409
Db 125 TCAGTCTCTGGGTCTCTATGTTCTCCGCGCAGCCCGCAGCACTTTCACAGCTTCTCTC 184
QY 410 ThrAlaCysValGlnThrProLeuGlyProAspTyrLysLysCysLeuIleArgGlnAsp 429
Db 185 ACTGCTGTGTCCAGAACCCCTTGTCTCTGATTATAAGGGCTGCTCATCGCGCGCAGC 244
QY 430 ProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeu 449
Db 245 CCCACAGGAACCTTCTTCTGTGTGGCTCAGCCGACCCACACAGCACTTTCGAGAGCTC 304
QY 450 LeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSer 469
Db 305 CTGGCAACCTGTGGATGGGGGCTGACGTAGATGGGTGGCGTGGCAGTCCCTCACTTC 364
QY 470 CysCysAlaProArgProLysGluLysSerAsnLeuIleValArgArgGlyCysAsn 489
Db 365 TGCTGTATCCCGACAGCCCAAGAAAGTCCAACTGATCGTGTCCAGAGAGTTCACAGC 424
QY 490 ProAlaProAlaProGlyCysSerPro---SerCysCysAlaLeuThrGlnLeuSerPhe 508
Db 425 CCACCCACATCATCTTGTGTTCCAGCCCAATCCCAATACCAGCTGAGTCAGATGACATT 484
QY 509 HisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThr 528
Db 485 CACAAGATCCCTGCTGACAGCTGGAGTGGCATGAGAACCTGGGCATCGGTCTCTCACC 544
QY 529 LysIlePheArgGlySerArgArgGluValValAspGlyGluThrHisAspSerGluVal 548
Db 545 AAGATTACCGGGGCTGTGCCATGAGGTGGTGGATGGGAGGCCGAAAGACAGAGGTG 604
QY 549 LeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAla 568
Db 605 CTGCTGAAGGTTCATGATGTCACCAAGCACAAAGAACTCATGAGTGCATTCTTGGAAACAGCG 664
QY 569 SerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAla 588
Db 665 AGCTTGATGAGCCAGTGTCTGACCGCATCTCGGTGCTGCTCCAGCGGTGTGCTGCTG 724
QY 589 GlyAspSerIleMetValGlnIlePheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608
Db 725 GGAGACACCATGTCAGGAATTTGTACCTTGGGGGCCATAGACATGATGATCTGCGA 784
QY 609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628
Db 785 AAACGTGGCCACCTGGTGCCAGCCAGCTGGAGTGGAGTGGTGGTGGTCAACACAGCTGCCTAC 844
QY 629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648
Db 845 GCCCTCAACTATCTGGAGGACAAAGGCTGCCCCATGCAATGCTCTCTGCCCGGAGGTG 904
QY 649 LeuLeuAlaArgGluGlyArgAspGlyAsnProProPheIleLysLeuSerAspProGly 668

```

```

905 CTCCTGGCTCGGAGGGGCTGATGGAGCGCCCTTCATCAAGCTGATGACCCCTGGG 964
669 ValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProThrValAlaPro 698
965 GTCAGCCCGCTGTGTAGCCTGGAGATGCTCACCGACAGGATCCCTGGTGGGCCCC 1024
689 GluCysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysThrGlyPheGlyAla 708
1025 GAGTGTCTCGGAGGCGCAGACACTTAGCTTGGAAAGCTGACAAAGTGGGCTTCGGCGCC 1084
709 ThrThrTropGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLys 728
1085 ACGGTCTGGGAAGTGTAGTGGCGTCACCATGCCATCAGTCCCTGGATCTGCTAAG 1144
729 LysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysThrThrGluLeuAla 748
1145 AAACCTCCAAATTTATGAGGACCGCAGCAGCTGCGCGCCCAAGTGGACAGAGCTGGCC 1204
749 GlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaIle 768
1205 CTGCTGATTCACAGTGCATGGCTATGAGCCGGTCCAGAGGCCCTCTTCCGAGCCGTC 1264
769 LeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrPro 788
1265 ATTCTGACCTCAATAGCTCATCTTTCAGACTATGAGCTCTCTCAGACCCACACCT 1324
789 GlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGln 808
1325 GGTGCCCTGGCACCTCGTGTGCTGGCTGGTGG---AATGGTGCCAGCTCTATGCTGCCAA 1381
809 AspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsn 828
1382 GACCCCAAGATCTTCGAGGAGAGACCTCAAGTACATCTCACAGCTGGGCAAGGGCAAC 1441
829 PheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuVal 848
1442 TTTGGCAGCTGGAGCTGTGCGCTATGACCCGCTAGCCCAATACAGGTGCCCTGGT 1501
849 AlaValLysGlnLeuGlnHisSerValProAspGlnArgAspPheGlnArgGluIle 868
1502 GCCGTGAACAGCTGCAGCAGCGGGCCAGACCAGCAGAGGGACTTTCAGCGGGAGATT 1561
869 GlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGly 888
1562 CAGATCTCTCAAGCACTGCAGTGAATTCATGTGCAAGTATCTGCTGCTCAGTATGCG 1621
889 ProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp 908
1622 CCGGGCCGCCAGAGCTGGGCTGCTCATGGAGTACCTGCGCCAGCGGCTGCTTGGCGCAC 1681
909 LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGln 927
1682 TTCCTGACGCGCACCGCGCGCGCTCGATGCGCCGCTCTCTCTATTCCTCGCAG 1741
928 IleCysLysGlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAla 947
1742 ATCTCAAGGGCATGGAGTACTTGGCTCCCGCCCTGCGTGCCAGCGGACCTGGCCGCC 1801
948 ArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLys 967
1802 CGAATCATCTCTGGAGAGCGGACACAGCTCAAGATCGCTGACTTGGCCCTAGCTAAG 1861
968 LeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePhe 987
1862 CTGCTGGCGCTTGACAAAGACTACTAGTGGTCCCGGAGCCAGGCGACGCCCATTTTC 1921
988 TrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSer 1007
1922 TGGTATGCCCCGAAATCCCTCTCGGACACATCTTCTCTCGCAGTCAGCTCTGGAGC 1981
1008 PheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027
1982 TTCGGGGTCTCTGTACGAGCTCTTCACTACTGCGACAAAAGCTGCGAGCCCTCGGCC 2041
1028 GiuPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGlu 1047
2042 GAGTTCCTCGGATGATGGATGTGAGCGGATGTCCCGCCCTCTCGCCCTCTTGGAA 2101
1048 LeuLeuAlaGluGlyArgArgLeuProProProProProProThrCysProThrGluValGlnGlu 1067
2102 CTGCTGAGGAGGCGCAGAGCTGCGCGGCGCTCTCTGCTGCTGCTGCTGAGGTTCAGGAG 2161
1068 LeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSer 1087
2162 CTATGAAGCTGTGCTGGGCGCCCTAGCCACAGAGCGGCGCATCATTCAGCGCCCTGGGC 2221
1088 ProGlnLeuAspProLeuTyrArgGlyArgProGly 1099
2222 CCCAGCTGACATGCTGTGGAGCGGAAGCGCGGG 2257
RESULT 15
HSU70065 13562 bp DNA linear PRI 27-JAN-1997
LOCUS Human JAK3 gene, complete cds.
DEFINITION U70065
ACCESSION U70065
VERSION U70065.1 GI:1800224
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 13562)
Riedy,M.C., Dutra,A.S., Blake,T.B., Modi,W., Lal,B.K., Davis,J.,
Bosse,A., O'Shea,J.J. and Johnston,J.A.
Genomic sequence, organization, and chromosomal localization of
human JAK3
Genomics 37 (1), 57-61 (1996)
JOURNAL 97079659
MEDLINE 8921370
PUBMED 2 (bases 1 to 13562)
REFERENCE Riedy,M.C., Dutra,A.S., Blake,T.B., Modi,W., Lal,B.K., Davis,J.,
AUTHORS Bosse,A., O'Shea,J.J. and Johnston,J.A.
Direct Submission
TITLE Submitted (09-SEP-1996) NIAHS, National Institutes of Health, 9000
JOURNAL Rockville Pike, Bethesda, MD 20892, USA
FEATURES
source
1. 13562
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.1"
201. 13562
/gene="JAK3"
join(201. 384,719. 842,1141. 1255,1462. 1607,2025. 2317,
2865. 2987,3082. 3350,4264. 4450,6303. 6430,6630. 6761,
7418. 7503,7918. 8044,8808. 8939,9019. 9171,9302. 9452,
10839. 10978,11059. 11256,12009. 12125,12400. 12969)
/gene="JAK3"
/codon_start=1
/product="JAK3"
/protein_id="AAC50950.1"
/db_xref="GI:1800225"
/translation="MAPSEETLIPQRSCLLSTEAGALHVLIPARGPPQRLSFS
SGDHLAEDLCQAAKAGSILFVHSLFALATEDLSCWFFPHSHIISVEDASTQVLLYRI
RSFYFNWFLGKCHRFGLRDLASAILDLPLVHLFAQHSRLVSLVGLSLUKEQ
GECLSLAVLDLARMAEQAPGELLKTVSKYACLPSPSLDLIOQLSFVTRAIRTV
PRALPRVAACQADBSLWAKYIMDLERLDPAGAETHVGLPGALGGHGLGLFRVAG
DGGIANTGEGVLQPPCFDIPRIVDISIKQAPRVGPAHRLVTVTRTNOILEASFPG
LPEALSVFALVDGFRLLTDSOHFKEKVAERPRLEVEAEQCHGITLDFALNKLKTG
GSRPGSVYLRSPQDFLLTVCVQNPGLKGLLRSPETGLRSTGLFLVLGLSRPSSL
RELLATCDWGLDGHVAVLTSCCIPRKEKSNLIVVQRHSPPTSSLVQSQQL
SQMTFHKIPADSLHEWHNLGHGFSFTKIRGCRHEVVDGEARTEVLLKVMADKRNCM
ESFLFAASMSQVSVRHLVLLHGVCMAGSTMVQFVHLGALDMLYLRKRGHLVPSLW
LQVVKQLAYALNYLEDKGLPGNVSARKVILLAREGADGSPFIIKLSDPGVSPVLSLE
MLTDKIPWPAECLREAGTSLSEADKNGFGATVWVFSGVTMPISALDPKAKLQFYED

```



QY	60	-----	60	QY	160	lyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaAa-rGluGlnAlaGlnA	180
Db	441	AGATTGGAAGCTGGAATAGTTGCTGCAAGATCAGATCGAGAGCTGGGGCTTTGGGGGA	500	Db	1520	GTGAGTGTCTCAGCTGGCGGTGTTGGACCTGGCCCGGATGGCGAGAGACAGCCGAC	1579
QY	60	-----	60	QY	180	rgProGlyGluLeuLeuLysThrVal	188
Db	501	TGAGTAGGAGTTTGTAAATGGAAGGGTGTGAGGGTTGCCTTCTGAGACAGGGAAAT	560	Db	1580	GGCCGGAGAGCTGTGTAAGACTGTCAAGGTGAGAGCCACAGGCTGTGGGAGCGGCTCT	1639
QY	60	-----	60	QY	188	-----	188
Db	561	GGCCTGTGACAGCGAGAGGTGTGACGGCACATGAAGGAAACAGCTGGTATACCTTGAG	620	Db	1640	GTTTGGAGTGAACAAGTGGGCTCCATCGGGGTTTGGGGGCTCCACCATGGAGTTC	1699
QY	60	-----	60	QY	188	-----	188
Db	621	GTATGGAAGATCTGGACGGTTGGGTATGATGTGGCAGCTTCCTGAAGGGCACAGATGGG	680	Db	1700	TCCTGCAAGCTTTCAGGGTGTTCCTATGACCCAGGGGCTCCACGAACCCAGGCTCTCC	1759
QY	61	-----	69	QY	188	-----	188
Db	681	TGACTCAGGAGGACCTGATGGGACCATCCCCCTGTAGGCATCTGCTGTGTACCACTC	740	Db	1760	ACCCCGAGCTCCAGCTGGACAGACACTGACCTCCCCAGCTAAAGGCTGTGGGCTC	1819
QY	69	rLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProProSerHisilePheCy	89	QY	188	-----	188
Db	741	CTCTTTGCTCTGGCCACGAGGACCTGTCTGTCTGCTGTTCCCGGAGCCACATCTTCTC	800	Db	1820	CCTGTCCGACCTGTGGCGGCAATGGCCCTCCCTCTCTGAGGTCCGGTCTCATACCT	1879
QY	89	sileGluAspValAspThrGlnValLeuValTyrArgLeuArg	103	QY	188	-----	188
Db	801	CGTGGAGATGCCAGCAGCCCAAGTCTGCTGTACAGGATTCCGGTACGCCCCCGCACAA	860	Db	1880	GACCTCTGAATGAGAGTCTGTGTGCTGTGCTGCTGCCCAACTAGGGCCGCAACCCAG	1939
QY	103	-----	103	QY	188	-----	188
Db	861	CGGATAGGACAGTTAGTATCAGGGTCTCTGTCATCTCTGTGGGACCCACACCCGAGCT	920	Db	1940	GGGCTAAAGCTGGGTTTGTGTATACCCCGGGGAGACCCCTCCCGACGCTGAGGACCG	1999
QY	103	-----	103	QY	189	-----	201
Db	921	CAGCGAGGTCTGTGTACTACTACCATTTTCATCAGCGTGCCAGCGAGGAGACAGTTGAG	980	Db	2000	GGTCCCTCCCTCCACCCCTGCAGCTACAGGCTCCCTACCCCAAGCTTCGCGACC	2059
QY	103	-----	103	QY	201	allleGlnGlyGlnAsnPheValThrArgArgGilleAa-rGluThrValValLeuAlaL	221
Db	981	GGTCCACCCAGCTCAGCACAAAGGCACACACAGAGGTTAAAGGTCAAGCTGGGCTCTT	1040	Db	2060	TGATCCAGGCTCGAGCTTCTGACGGGAGGGGTATTTCGAGGACGCTGCGCAGAGCC	2119
QY	103	-----	103	QY	221	euLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeu	240
Db	1041	TGAGGCTCCGCAAGTCTGCTCCCTGCGCCCCACCATAATGTCACTCTACTGAGGTGG	1100	Db	2120	TGCC-GCGCGTGGCGCTCCAGGACGCGGACCTCGCTCATGGCCAGTATCATG	2178
QY	104	-----	109	QY	241	AspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly	260
Db	1101	GTTGCACCTTTTCATCCAGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1160	Db	2179	GACCTGAGCGCTGGATCCAGCCGGGCGCCCGAGACCTTCCACGTGGGCTCCCTGGG	2238
QY	109	rpPheGlyLeuGluThrCysHisArgPheGlyLeuArgGlyLeuAspLeuThrSerAlaIleL	129	QY	261	Ala-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle-Pr	277
Db	1161	GGTTGGGCTGGAGAAGTGCACCGCTTCGGGCTACGCAAGGATTTGGCCAGTGTATTC	1220	Db	2239	GCCCTTGGTGGCCACGACGGGCTGGGCTGTTCCGCGTGGGTGGTGAGCGGCGCATCGCC	2298
QY	129	euAspLeuHisValLeuGluHisLeuPheAla-----	139	QY	277	o-----	277
Db	1221	TTGACCTGCCAGTCTTGAGCACCTCTTTTG-CCAGGTGGGGTCTCTGCTGGGGTTTGAC	1279	Db	2299	TGACCCAGGAGAGAACAGGTGAGACGGCTTCCCTCCCTGGCGGGGCCAAATGGGGGGG	2358
QY	139	-----	139	QY	277	-----	277
Db	1280	CAGGGGTTGGGGTCCAAATGGCAACATGAGGACTGGCAATGCAATCAGGTGGGGCTCG	1339	Db	2359	GTTTGGGAGGGCGGGAGATTGTAGGGATGTGGGCGGAGCCAAACGAAAGACTTG	2418
QY	139	-----	139	QY	277	-----	277
Db	1340	TCTGACCTCCCTGTGGCAGTCCAGGGGTGGGGGTGAGGATTTGGGGGGGTCTGCA	1399	Db	2419	GGGAAGTGGCGAGGCTTAATGAGGGCGGGGCTTAGTGAGGAGGAGACTGCGGGAATG	2478
QY	140	-----	140	QY	277	-----	277
Db	1400	GTTTAAACACAGGCTTGAAGTTGGGTGGCTTCAGCTGATGCTGCTGTGGCGGCCCTCC	1459	Db	2479	GGAGGGCAAACTGATGTAAGGGTTGGCTGAGCGACCCGGGAATGAAATGGGAGGGT	2538
QY	140	-----	140	QY	277	-----	277
Db	1460	AGCACCGCAGTACCTGGTGTAGTGGGCGCTCCCGTGGGCTCAGTCTCAAGAGGACAGG	1519	Db	2539	TGAGGAGTGGGTGGGCTGGAAACGAGGCAAGACTAAGAGGAGAGTGTAGGCTCGG	2598
QY	1460	AGCACCGCAGTACCTGGTGTAGTGGGCGCTCCCGTGGGCTCAGTCTCAAGAGGACAGG	1519	QY	277	-----	277

Db	2599	AGAGCTGGAGGCTGCGAAGCGGGGCGAGGATCAAGGGCGGGTCCGGCAGAGAGAGG	2658
QY	278	-----TtpSerSerAsnAspGluLeuPhe-----	285
Db	2659	AGCGGGGATTTGCAGAGCGGTGGGAGATGTTTGGGATGGTAGGGAGGTGTGAGA	2718
QY	285	-----	285
Db	2719	GGTTGGAGTTGATAGGAGGGGCGCGCTTGGAAGGTTGAATGGCAAGGATAGGAG	2778
QY	285	-----	285
Db	2779	TGGATGGTGTGGCTTGGGGTGGGTTGCATGGCGTGGTTTGGCGGACCCGCTGGGCC	2838
QY	286	-----GlnThrPheCysAspPheProGluLeuVa	295
Db	2839	CACTACGGTAANCCCCAACCAACCACTCTCCAGCCCTCTCGGACTTTCCAGAAATCGT	2898
QY	295	laspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValTh	315
Db	2899	AGACATTAGCATCAAGACAGCCCGCGGTGGCCCGCGGAGAGACACCGCTGTGTAC	2958
QY	315	rValThrArgMetAspGlyHisIleLeu-----	324
Db	2959	TGTTACAGGACAGACAACCAAGATTTAGGTGTGCAGGATTCCTCCCTTCAGCCT	3018
QY	324	-----	324
Db	3019	TACCCGAGGGCGGACCGGACCCCTCGGTTTCACTGGGCTCTGACGCTGTGCCTCG	3078
QY	325	--GluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyT	344
Db	3079	CAGAGCCGAGTTCCAGGCTGCCAGGCTCTGTCTGTGGCGCTCTGGACGGCT	3138
QY	344	yrPheArgLeuIleCysAspSerArgHisTyPheCysLysGluValAlaProProArgL	364
Db	3139	ACTTCGGGTGACCAAGGACTCCAGCACTTCTTGCAAGAGGTGGCACCGCGGAGC	3198
QY	364	euLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisL	384
Db	3199	TGCTGGAGGAGTGGCGGAGTAGTCCACGGCCCCATCACTCTGGACTTTGGCATCAACA	3258
QY	384	ysLeuLysAlaAlaGlySerLeuProGlyThrTyIleLeuArgArgSerProGlnAspT	404
Db	3259	AGCTCAAGACTGGGGCTCACGTCCTCGCTCTATGTTCTCGCGCGACGCCCCAGACT	3318
QY	404	yrAspSerPheLeuLeuThrAlaCysValGln-Thr-----	415
Db	3319	TTGACAGCTTCCTCCTCACTGTCTGTGTCCAGGTGGTCTACTGTAGGGTGGGTAGTGG	3378
QY	415	-----	415
Db	3379	AGGGCTGCCTGGACAAGAGGTTACCTCTGAATTGAGATTTACTCAGATCTTACTCAGCAT	3438
QY	415	-----	415
Db	3439	TTGGTTCTCGAAACAAATAGGGAACCGACACCCCGGTTAAGCGAACACGCTATTCT	3498
QY	415	-----	415
Db	3499	CTGGTCTAAATACTGGAATTAGTTGGTGGTCTTCAGGATAAAGACTGAAGCTGGGCG	3558
QY	415	-----	415
Db	3559	TGTCGGTCCCCCTGTTATCCAGCCCTTTGGAGGCTGACGTAAGCATATCCCGAGATC	3618
QY	415	-----	415
Db	3619	GGGATATCCACCCATCTGGTCCAAATGGTTAAACCCCTGTCCACCTGTAATATGTAA	3678
QY	415	-----	415

Db	3679	CTCAGCTGGGCATGGTAATGGGCATCTGTTAATCCACCTACTCTGGGAAGCTCAAGGCAT	3738
QY	415	-----	415
Db	3739	GTTTTCTAAGGTTTTTAAAAAAGGGGGGGGTTTTTTTGGGGTTTTTGGGGAAAAAAA	3798
QY	415	-----	415
Db	3799	AAGGGGTGGGGNAAAAAAGGGGTAAAAACCTTTTTTTAGTCTTCGAGAGCTCTCTGGCAAC	3858
QY	415	-----	415
Db	3859	CTGCTGGGATGGGGGCTGCACGTAGATGGGTGGCAGTGACCTCCTCCTCTCTGTGTAC	3918
QY	415	-----	415
Db	3919	CCCAAAATTTGGGGGTGGGTGGCCAGTTTAAAGGTTTAAAGGTTTGGNAACCTTCGGG	3978
QY	415	-----	415
Db	3979	GCCCAAGGGGAACCCCTTTTTTTTTTAAAAAATTTTAAATAATTTTGGGGGGGGTAAGG	4038
QY	415	-----	415
Db	4039	GGGATTTTTTTTCCCAAAATTTGGGGGGGGGGCGGGGCATTTTTTTTAAACCCGGGG	4098
QY	415	-----	415
Db	4099	GGGGGGGTTCCAGGGGGCCAGGATAAACCTTTAAATTTCCCGGGGAAAAAGGCCCA	4158
QY	415	-----	415
Db	4159	AATTCTTTTCCCAAAAAAAGAAAAAAGAAAGAAAGAAAGAAAGAAAGAGAG	4218
QY	416	-----ProLeuGlyPro	419
Db	4219	AGTGTCTGTGCTCAGTCCCACTCAGGGGCCACTCTTCTTTGCAGAACCCCTTGTCT	4278
QY	420	AspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeu	439
Db	4279	GATTAAAGGCTGCTCATCGGGCGAGCCCAAGGAAACCTTCTCTGTGTGGCCCTC	4338
QY	440	SerGlnProHisArgSerLeuArgGluLeuLeuAlaCysGlyTyrAsnSerGlyLeuArg	459
Db	4339	AGCCAGCCCAACAGCAGTCTTCGAGAGCTCTCGGCAACCTCTGTGGATGGGGGCTGCAC	4398
QY	460	ValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArg-ProLys-----	476
Db	4399	GTAGATGGGTGGCAGTGACCTCACTTCTCTGTATCCCCCAGACCCCAAGGTGAGGCC	4458
QY	476	-----	476
Db	4459	CCTGGTCCCTCGAATGAGTGGCTGATCTGGGACCTGGTTTCTCTATGTCTGATAACAG	4518
QY	476	-----	476
Db	4519	CTCCTGTGTGATGGCAAGTGGTAGAAACTGCAGATCATGTGGGTAGGGAAGAAAGG	4578
QY	476	-----	476
Db	4579	CGATTTGTCTTCAGGAAGTTAGAGCCATCTAACCCCTAGGGCTGGATTGCCACGGTA	4638
QY	476	-----	476
Db	4639	CAGCTAGATGTAGCCATGTACCTGGACCTTGNTTTATCTCCAGAAATCTCAGTCTTTC	4698
QY	476	-----	476
Db	4699	TTCTGGGTGACCTTCTACTCTCTGGCACAAACTCCCCCTGTATGGGGAACAAATTA	4758
QY	476	-----	476
Db	4759	GAGTATGTTGGAGCCCTTTCTCTGAGTTCCTTCCAGCAATTCCTCGGTGAAGAAACCCGAT	4818



```
QY 476 ----- 476
Db 5899 GTGGTTCTCTCGATAAAACACAGAAATTTCCCGGGGAAACCCGGGAATTTTTT 5958
QY 476 ----- 476
Db 5959 TTTTAAAGAACAAAAAATAATAGCCGGGACCGGTGGCACACACCTGTAGTCCAGC 6018
QY 476 ----- 476
Db 6019 TATTCGGAGCCGAGGACAGAAATCGCTTGAACCTGGGAGGTGGAGTTGCAGTGAGC 6078
QY 476 ----- 476
Db 6079 TGAGATCGCACCATTGCCACCAGCCTTGATGACAGAGTGTAGATCCGTCTCAGACAG 6138
QY 476 ----- 476
Db 6139 CAGCAACAGCAAAACAAAAACAACAAAAAGCATGTGCTGAACTCTTATCTCAG 6198
QY 476 ----- 476
Db 6199 GGTGGGTTCTAGAGGTACCTCAAACTAAGCATGATTAGCTAACCTTGGGACTTT 6258
QY 477 ----- GluLysSerAsnLeu11 482
Db 6259 TCACCTCTGATTCTGTGTTTCTCCCTCATCTCTCCCATAGAAAGTCCAACTGT 6318
QY 482 eValValatArgGlyCysAsnProAlaProAlaProGlyCysSerPro--SerCysCy 501
Db 6319 CGTGGTCCAGAGAGTACAGCCCAACCATCATCTCTGGTTCAGCCCAATCCCAATA 6378
QY 501 sAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrp 518
Db 6379 CCAGCTGAGTCAGATGACATTTCAAGAATCCCTGCTGACAGCCTGGAGTGGGTAAGAG 6438
QY 518 ----- 518
Db 6439 CCCTGGGAAATAGGCGCATACCTCAGTCTGGGGTCCAGAGACTCAGATCGGTGCGCTCAG 6498
QY 518 ----- 518
Db 6499 GCATATGCTATAATTTTACCTTGCTCGCTTTCCTCATCTGTAAATGGGCGCAGACT 6558
QY 518 ----- 518
Db 6559 ATGTCTCGCTGGGCTGGATCCTGAGGAACCCCTCACTGGCTCTTCTGTGTCCTCC 6618
QY 519 ----- HisGluAsnLeuGlyHisGlySerPheThrLysIlePheHargGlySerA 535
Db 6619 TCACCATTCAGCATGAGAACCTGGGCTATGGGTCTTCCCAAGATTTACCGGGGCTGTC 6678
QY 535 rgArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAsps 555
Db 6679 GCCATGAGGTGGTGGAGGGGCGCCGAAAGACAGAAAGTCTGCTGAGAGTCTAGGATG 6738
QY 555 erArgHisArgAsnCysMetGluSer 563
Db 6739 CCAAGCACAGAACTGCATGGAGGT-GAGAGGCAATGTGGACAGACTTTTGGAGTCCGG 6797
QY 563 ----- 563
Db 6798 GGCTGGCTTGAGCGGGGCTCGTGGATGCAGAGACATTTACAACACACAGGGAACCTGGC 6857
QY 563 ----- 563
Db 6858 GTGGTGGCTGATGCTGCATCATAGCGGGACTCTGCTTGCACAAAAAATTTTAGGCC 6917
QY 563 ----- 563
Db 6918 GGGCGGCTGGCTCAGCCTGTGAATCCAGCAGCTTTGGAGAGCTTGGAGGAGAGAGATCA 6977
QY 563 ----- 563
```



Db	6978	TCGAGGTGGGAGTCTCGAGACCAACCTGCCCAACATGAGAAACCCCGCTCTACTAA	7037	Db	8058	CCCCAGGGGATTGGTTTGGGGGAGGGGGGTTTAAAAAACCCAAACCCCTAAAT	8117
QY	563	-----	563	QY	632	-----	632
Db	7038	AAAAACCAAAATTAACCGGGCGTGGTGGTGATGCTGTAAATCCACAGCTACTCGGAGTC	7097	Db	8118	TTTGGGGGGGGTGTGGGCTGTTCCTCCCTCCATTCCCGGAAAGTTTITTTTTTTT	8177
QY	563	-----	563	QY	632	-----	632
Db	7098	TGAGGCAGGAGAAATTCCTTGAATCTGGGAAGCGGAGTTCACAGTGAGTCAAGATCGGCC	7157	Db	8178	AAAGGGGAAAAGGAATGGTTTITTTTTTAAAGATGTGTAAAGTTTGGGGGGGG	8237
QY	563	-----	563	QY	632	-----	632
Db	7158	ACTGCATCCAGCTGGGTGACAGAGGAGACTCTGTCTCAAAAAACAANCAANCAAAA	7217	Db	8238	CAITTTAGTTCTTTTTCGGAAAAATGGNAGAAAAATTTTTTTTTTTTTTTTNGGTT	8297
QY	563	-----	563	QY	632	-----	632
Db	7218	CAAAAAACCATAGATGATAGTGGAACTTCTGTCCCGTATCAGAAAAATCATGGTAGTGT	7277	Db	8298	TTGGGAAAAGGGGTGGCCGGGAAGAGGTTTCTTTTTTGGAAAAAGGACATGGCC	8357
QY	563	-----	563	QY	632	-----	632
Db	7278	GTGTGCACATAATGGCAGACTCCAGGGCCAAAGGTGACCTGTGCCAGGTGTCCCTTAAG	7337	Db	8358	CTTCACCTGGAGTGGTTTTCATGAAAAAGAGGAGGATTCCTCCCTGGGCCCTATCAGCATT	8417
QY	563	-----	563	QY	632	-----	632
Db	7338	GCAGTCTGTGAGCAAAAAATTTGGGATTATTGGAGTGAAGAAACCCACGCATCTTCTC	7397	Db	8418	ACCAAAAGCAAGCCCGCGGAGAAAGGGAGTATTTCCCTGGGCCCTATCAGCATT	8477
QY	564	-----	564	QY	632	-----	632
Db	7398	TCCCTTCCACCTTCCCGAGTCATCTCTGGAAGCAGCGAGCTTGATGAGCAAGTGTCTG	7457	Db	8478	TCCGTGCCAGGAGACAGAGGAAAAAGTGGGAAGGGAATTAATTGGAAGCTTGGAGCAG	8537
QY	576	-----	590	QY	632	-----	632
Db	7458	ACCGGCATCTGTGTCTCTCCAGCGCTGTGCATGCTGGAGACAGTGTGAGTCTCTAG	7517	Db	8538	GGCCCTGGAAATCCAGGTGAGGAGTTAAATTTTTTTTCAAAGGCATTGGGGTTAACAG	8597
QY	590	-----	590	QY	632	-----	632
Db	7518	GGGGTGTAAAGGGGAGAGGAGGAGAAAGAGGCGAGGGGGAAGAGAGTAGGGGCA	7577	Db	8598	ATGGGTGTGTGTGTGAGAGGGGAGGACATACATGTCATTTTTTGGGGGTGAAGTAGAG	8657
QY	590	-----	590	QY	632	-----	632
Db	7578	GAGAGGAGAGAAAGCCGGGAGAGAGCCAAACGGGAGCAACCAAGAGAGGTTAGCGAGA	7637	Db	8658	GGGCAAGACAGGATATCAGGATTTCCAGGACGATGTGTGGAGCATGTTTGGAGCAGTACCA	8717
QY	590	-----	590	QY	632	-----	632
Db	7638	GGATAAAAACTGGCTGACGTTTATAAAGATAGGGGGAGAGCCCTGCCTCGAGGC	7697	Db	8718	GTGGTTTGAAGGATATAGGAGTTTCCAAACAGACATCTTATCAAACTCC	8777
QY	590	-----	590	QY	633	-----	643
Db	7698	AAATTTATAGCCCGGAGGAGAGCTTCTGGCTTGGGGGGGGCTGAGAGGAGACCC	7757	Db	8778	CGGGCATTTTCGTGTGTGTGGCCCCCTTAGAGGACAAAGGCTGCCCATGGCAATGTC	8837
QY	590	-----	590	QY	644	SerAlaArgLysValLeuLeuAlaArgGluGlyAspGlyAsnProPheIleLys	663
Db	7758	GCCCCTTTAAACAAACTAGAAAGACCTCTGGGGATAATCCCCCGCTTATTATAACCTGG	7817	Db	8838	TCTGCCGAAAGGTCTCTCTGGCTCGGAGGGGGCTGATGGAGCCGCCCTTCATCAAG	8897
QY	590	-----	590	QY	664	LeuSerAspProGlyValSerProThrValLeuSerLeuGlu	677
Db	7818	AGAGAACTCAACGAAAAAGAGATCCCTGGGATGCCCCAGATATATTTCCGGAGTGGC	7877	Db	8898	CTAGTGACCTGGGGTCAAGCCCGCTGTGTAAAGCTGGAGTAAGTTCTCTGGAGGTGG	8957
QY	591	-----	591	QY	677	-----	677
Db	7878	AAAGAACTAAACTTTGGAATCTAGCAAAATATGGCTGAGCACCATTGGTGAGGAATTTG	7937	Db	8958	AGGAGGGGGGCTCAACAGAGGCAAGGAGTGGATCCCTGATCCCACTTTCATTCCTCA	9017
QY	598	alTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerT	618	QY	678	-MetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLe	697
Db	7938	TACACTGGGGCCCATAGCATGTATCTGCAAAACGTGGCCACCTGGTGGCAGCCAGCT	7997	Db	9018	GATGCTCACCACAGGATCTCCCTGGTGGCCCCCAGAGTGTCTCCGGAGGGGCGACACT	9077
QY	618	rpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyr	632	QY	697	uCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgG	717
Db	7998	GGAAGCTGAGTGTCTCAACAGCTGGCTAGCCCTCAACTATCTGTGTAGTGTCTCTC	8057	Db	9078	TAGCTTGAAGCTCAAGTGGGCTTCGGGCCACCGCTCTGGGAAGTGTTTAGTGGCGT	9137
QY	632	-----	632	QY	717	yProAlaHisIleThrSerLeuGluProAlaLys	728
				Db	9138	CACCATGCCCATAGTGGCTGGATCTGTAGGTGAGGCCCTCACCCGGCATCGGT	9197

QY	728	-----	728	QY	778	-----	778
Db	9198	CTCCGAACCCCACTTTGACAGCGGGGAGACTGACATCCAGTCTGGGGAGATTGGGGTG	9257	Db	10278	TCCACACAAAGAACCCAAAAAGCAGCTAGCAGATCATTTCAAAGTGTGAAATGTGAT	10337
QY	729	-----	734	QY	778	-----	778
Db	9258	GGTCTATTGGTTGGGATTACCGACTGCTCCTCTCACCTCAGAACTCCAATTATATG	9317	Db	10338	CTTTCTCAATTTTGGGCTGTTCCCTCTTCAAAATTCATCTTTCAAAGGGCATTTTG	10397
QY	734	luAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnC	754	QY	778	-----	778
Db	9318	AGACCGGACGACGCTGCCGCGCCCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGT	9377	Db	10398	AAGAGTATTACGGTAGAGGAGGTGTCTCTATCTCGTCCAACTATGGGCAATGGCA	10457
QY	754	ysMetAlaTyAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnG	774	QY	778	-----	778
Db	9378	GCATGCGCTATGAGCGGTCAGAGGCCCTCTCCGAGCGCTCAITTCGTGACCTCAATA	9437	Db	10458	GAGTGACCGTGAATATATAGGCAITTTTGTATCGAGTCGACGCCCTATAGTCGTATAG	10517
QY	774	lyLeulleThrSer-----	778	QY	778	-----	778
Db	9438	GCCTCATCTCTTCAGGTGCCCGCTGGACGGTTGGTGGGGGCTGTGATGTCATAT	9497	Db	10518	AGCTTAGTATATAGAACCGCAGGAAATTTGATGTTCTTCAATTTATGCCCCCATCC	10577
QY	778	-----	778	QY	778	-----	778
Db	9498	TGGGCCAGTGAAGGAGCGTGTTCAGTAGGACCGCCTGTGTCTGTGTAGGTT	9557	Db	10578	TTTTGTAAACCAATTTGGGAAGTTGGCACCAGGTTTTCACCTGCACGAGGAGGCC	10637
QY	778	-----	778	QY	778	-----	778
Db	9558	GGAGGGTGTATGACTATAGACAGACAGTGTGAGCTTCAACAGCTACTGTAAACGAACATC	9617	Db	10638	TTCTTTTCCCTGGAACAAGAGTGGGAAGACACCTTTTTTCCAGTCTCTCCACAAGT	10697
QY	778	-----	778	QY	778	-----	778
Db	9618	CCCTCAGTTCAGGTTCAACATGATGATATTTATTTATTTATTTATTTTGGAGACAGAG	9677	Db	10698	ACCTTTTGGCTTGCAATGGCCCTGCCATATATGCACAGAGGGTCAATATGCAGATGGA	10757
QY	778	-----	778	QY	778	-----	778
Db	9678	TCTTGCTCTGTGCTAGGCTGGAGTGCAGTGGCACAAATTTGGCTCCTCAGAACCTCTA	9737	Db	10758	GGTTGCACCAAGTCAAACTCAGAGTGGGACCCAGAGTGAGAGGGCTGCTTACCACCT	10817
QY	778	-----	778	QY	779	-----	792
Db	9738	CCTCCTGAACCTCTAGTTCAAGGATTTCTCCTGCTCAGACTCCTGAGTAGCTGGGNTT	9797	Db	10818	GCCCATGCCCCACCCACCACTATGAGTCTCTCAGACCCACACCTGGTGCCTGG	10877
QY	778	-----	778	QY	792	erProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyAlaCysGlnAspProAlaI	812
Db	9798	ACAGTGTGTGCCACCACGCCAGCTAATTTTGGTTTGGTGGGAGATTTTGGGGT	9857	Db	10878	CACCTCGTGTGTCTGTGG---AATGGTCCCCAGCTCTATGCTGTCCAGACCCACGA	10934
QY	778	-----	778	QY	812	lePheGluAlaArgHisLeuLysTyrlleSerLeuLeuLys-----	826
Db	9858	TTCAACCATATTTGGCTAGGGTGTCTTCGAACCTCTTGACCCCAAGGTGATCCGCCCCAC	9917	Db	10935	TCTTCGAGGAGACACCTCAAGTACATCTCACAGCTGGGCAA-GGTAAAGGTGGCAGGG	10993
QY	778	-----	778	QY	826	-----	826
Db	9918	CTCTACCTCCTAAAGTCTGGGATTCAAATTTCTCGTACTTGGCTTGTGTTGATATAT	9977	Db	10994	CCAGGTGGTGGAGAGGCGAGCATCCAGTACCTGGACATCAGTCCCGCTATCC	11053
QY	778	-----	778	QY	827	-----	845
Db	9978	TGCCAAATTTAGTCGAGATATAGTTTGGCAGTAGACTCATGATTTTCAGTTCTGTGAG	10037	Db	11054	TCGAGGGCACTTTGGCAGCGTGGAGCTGTCGGCTATGACCCGCTAGCGCAATACAG	11113
QY	778	-----	778	QY	845	lyProLeuValAlaVallysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheG	865
Db	10038	ACTTTGTAAACAAGAGTTTCAAGTGTCAAAGGAGGAACCGGATTTGTTTATAATAAA	10097	Db	11114	GTGCCCTGTGGCCGTGAAACAGCTGCAGCAGCGGGCCAGACCAGCAGAGGAGCTTTC	11173
QY	778	-----	778	QY	865	lnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVallysTyArgGlyV	885
Db	10098	CCTGCAGTATTAAATAGGAATCATAATAATGACTCAGTATATACCAGGAGGAACAC	10157	Db	11174	AGCGGAGATTTCAGATCCTCAAAGCACTGCAGTGATTTTATTGTCAGTATCGTGTG	11233
QY	778	-----	778	QY	885	alSerTyGly-----	888
Db	10158	CCAAATAAGTTAAGTTGCTTAAAGAAAGAAAGATGGGTGATGGTAAATAATAATTA	10217	Db	11234	TCAGCTATGGCCCGGTGAGCCAGCTCCGGATGAGTGAACCAAGACGTATGGTCTTT	11293
QY	778	-----	778	QY	888	-----	888
Db	10218	AAGTCCAGCCAAACAAATTAAGCGTATAAGATTCTGTAAACTGAAACATAAGCAAAA	10277	Db	11294	TCAAAGTGCACATTTTACCCTCTGCGAGGCCACTTTAGTAGGTGGAAACGTGATTA	11353
QY	778	-----	778	QY	888	-----	888





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 00:42:56 ; Search time 798.095 seconds  
(without alignments)  
3717.204 Million cell updates/sec

Title: US-09-397-967A-16  
Perfect score: 5860  
Sequence: 1 MAPSEETPLIPQRSCSLSS.....RPAFATLSPQLDLWRGPG 1099

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO.spool/US09397967/runat\_03022004.175631.28568/app\_query.fasta.1.1486  
-DB=N\_Geneseq\_19Jun03 -QMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LCODEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=ext -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEP SIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967.csgn.1.1.501 @runat\_03022004.175631.28568 -NCFU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5223	89.1	4016	17	AAT11083	Mouse JAK3 protein
2	4977	78.4	3620	25	ACA56457	Human signalling p
3	4418	75.4	3807	17	AAT30862	Protein tyrosine k
4	2629	44.9	3435	19	RAV61801	JAK2 protein encod
5	2627	44.8	6037	25	ACC46252	Human dithp intrac
6	2621	44.7	4482	20	AAX80971	Human JAK2 kinase
7	2621	44.7	4482	21	PAZ58947	Human JAK2 kinase
8	2621	44.7	5117	24	AAD24311	Human JAK2 kinase
9	2619.5	44.7	3629	16	AAQ85412	Murine JAK2 kinase
10	2619.5	44.7	3629	21	AAC66244	Murine JAK2 kinase
11	2619.5	44.7	3629	22	AAD03607	JAK2 polynucleotid
12	2588	44.2	3503	21	ABQ76194	Murine (Janus kina
13	2486.5	42.4	3900	25	ACA55699	Human JAK2 CDNA.
14	2376.5	40.6	3473	13	AAQ5307	Human signalling p
15	2360	40.3	3495	24	AAD22680	JAK2 encoding DNA.
16	1908.5	32.6	4191	24	ABK72331	Murine JAK2 tyrosi
17	1904	32.5	3429	16	AAQ85413	DNA encoding lymph
18	1904	32.5	3429	21	AAC66245	Human JAK1 kinase
19	1904	32.5	3429	22	AAD03608	Human JAK1 kinase
20	1904	32.5	3540	13	AAQ26039	Human (Janus kina
21	1904	32.5	3541	24	ABK84065	JAK1 encoding DNA.
22	1904	32.5	3541	24	ABN95881	Human CDNA differe
23	1904	32.5	3541	24	ABK72301	Gene #2379 used to
24	1904	32.5	3541	24	ABL61826	Lymphona associate
25	1842	31.4	3538	24	ABL72332	Colon adenocarcino
26	1817.5	31.0	4080	25	ACA56755	DNA encoding lymph
27	1817.5	31.0	4176	24	ABK84572	Human signalling p
28	1817.5	31.0	4176	24	AAD24312	Human CDNA differe
29	1815.5	31.0	3561	16	AAQ85414	Human TYK2 kinase
30	1815.5	31.0	3561	21	AAC66246	Human TYK2 kinase
31	1815.5	31.0	3561	22	AAD03609	Human TYK2 kinase
32	1813.5	30.9	4248	22	AAH32035	Human polynucleoti
33	1806.5	30.8	4218	22	AAH98351	Human EST-derived
34	1806.5	30.8	4220	22	AAH53019	Human polynucleoti
35	1564	26.7	4294	23	AAH85150	DNA encoding novel
36	1010	17.2	795	22	AAH10808	Human Janus kinase
37	948.5	16.2	726	22	AAH10807	Human Janus kinase
38	857.5	14.6	778	22	AAH10809	Human Janus kinase
39	816.5	13.9	2000	22	AAK78097	Human immune/haema
40	746	12.7	2327	20	AAZ33632	Human breast tumou
41	713	12.2	2338	21	AAF18071	Lung cancer associ
42	680.5	11.6	5045	23	ABL02929	Drosophila melanog
43	630	10.8	1026	23	AAH85149	DNA encoding novel
44	592.5	10.1	811	22	AAH10806	Human Janus kinase
45	507	8.7	3222	22	AAK78095	Human immune/haema

## ALIGNMENTS

RESULT 1  
AAT11083  
ID AAT11083 standard; CDNA; 4016 BP.  
XX  
AC AAT11083;  
XX  
XX 08-APR-1996 (first entry)  
XX  
XX Mouse JAK3 protein-tyrosine-kinase gene.  
XX  
XX Mouse; Janus kinase; JAK3; protein-tyrosine-kinase; cytokine;  
KW signal transduction; 32Dcl3; reverse transcription; PCR;  
KW polymerase chain reaction; primer; interleukin-3;  
KW granulocyte-macrophage colony stimulating factor; cDNA probe;  
KW granulocyte colony stimulating factor; Escherichia coli;  
KW haematopoietic cell; differentiation; leukaemia; therapy; ss.

OS Mus musculus.  
 XX Key Location/Qualifiers  
 FT CDS 7..3906  
 FT /\*tag= a  
 FT /product= JAK3 protein-tyrosine-kinase  
 FT /note= "BC-2.7.1.112"  
 XX  
 PN W09601838-AL.  
 XX  
 XX 25-JAN-1996. 95WO-US08354.  
 XX  
 XX 28-JUN-1995;  
 XX  
 XX 08-JUL-1994; 94US-0272368.  
 XX  
 XX (UTEM ) UNIV TEMPLE.  
 XX  
 XX Rane SG, Reddy PE;  
 XX  
 XX WPI; 1996-097584/10.  
 XX P-PSDB; AAR88560.  
 XX  
 XX JAK3 protein tyrosine kinase and DNA encoding it - useful to induce  
 PT differentiation of haematopoietic cells to treat or control  
 PT leukaemia  
 XX  
 XX Claim 4; Page 26-28; 49pp; English.  
 XX  
 XX The sequence encodes a novel protein-tyrosine-kinase of the Janus  
 CC kinase (JAK) family (involved in cytokine signal transduction),  
 CC designated JAK3. The sequence has been isolated from mouse 32Dcl3  
 CC cells growing in interleukin-3 or granulocyte-macrophage colony  
 CC stimulating factor by reverse transcription-polymerase chain  
 CC reaction amplification, using primers J3 and J4 (AAR11084-T11085),  
 CC based on conserved sequences in JAK JH2 and JH1 domains. A 950-bp  
 CC cDNA probe is isolated, which is then used to screen a granulocyte  
 CC colony stimulating factor-treated 32Dcl3 cDNA library in phage  
 CC lambda-gt10 in Escherichia coli to isolate the full-length gene.  
 CC The JAK3 protein is useful for stimulating the differentiation of  
 CC haematopoietic cells in therapy or control of leukaemia.  
 XX  
 XX Sequence 4016 BP; 742 A; 1203 C; 1235 G; 836 T; 0 other;  
 SQ

Alignment Scores:  
 Pred. No.: 0 Length: 4016  
 Score: 5223.00 Matches: 1060  
 Percent Similarity: 80.59% Conservative: 3  
 Best Local Similarity: 80.36% Mismatches: 32  
 Query Match: 89.13% Indels: 235  
 DB: 17 Gaps: 6

US-09-397-967A-16 (1-1099) x AAT11083 (1-4016)

QY 1 MetAlaProSerGluThrProLeuileProGlnArgSerCysSerLeuSerSer 20  
 DB 7 ATGGCACTCCAGTGGAGACACCTCTGATCCCTCAGCGCTCTTGACGCTCTCATCC 66  
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
 DB 67 TCAGAGGCAGAGCCCTGATGTGCTCTCTCCCGGGGACCTGGGCTCCGACGGA 126  
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaIysAla 60  
 DB 127 TTGTCAATCTCTTTGGGGACTACTTGCTGAGATTATGTGCGAGCTGCCAAGGCC 186  
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 DB 187 TGTGGATCTCGCTCTGTTTATCATTCCTGCTTTTGGCTCTGGCCACTGAGACTCTCTTGC 246  
 QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 DB 247 TGGTTTCCCGCCCAAGCCACATCTTCTGTCATAGGACGCTGGACACTCAAGTCTTGGTCTAC 306

QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCystHisArgPheGlyLeu 120  
 DB 307 AGGCTACGCTTTTATTTCCCTGACTGGTTTGGCTGGAGACATGTCCAGGCTTGGGCTG 366  
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
 DB 367 CGCAAGATTTCACCAAGTCCCATCTTGTGACTTACATGTTTAAACAATCTCTTTGCTCAG 426  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
 DB 427 CACCGCAGTGACCTGTGTGAGTGGGCGCTCCCGTGGGCTTAGCATGAAGGACAGGA 486  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
 DB 487 GAGTTCTTGAGCTGCGCTGCTGGACTTGGCCAGATGGCTCGTGAGCAGGCCAGCGC 546  
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
 DB 547 CCAGGAGAGCTGCTGAAGACGGTCAGTTACAAAGCCTGTCTGCCGCCAGCCTGCGGAT 606  
 QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
 DB 607 GTGATCCAGGCCAGAACTTCTGTCACAGCAGCGCATCCGCAAGCCGTGTCTTGGCG 666  
 QY 221 LeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeu 240  
 DB 667 CTGC--GCGTGTGTCGCTGCCAGCCGACCTACGGCTCATGGCCAAAGTATAT-CTG 723  
 QY 241 AspLeuGluArgLeuHisProAlaAlaThrGluThrPheArgValGlyLeuProGly 260  
 DB 724 GACCTGAGCGGCTACATCCAGCGGCCACACCGAGACCTTCCGTGTGGGGCTCCCGGGC 783  
 QY 261 AlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer 280  
 DB 784 GCCCAGAGGAGCGCGGGCTTCTGCTGTGGCGGGGACAAACGGCATCTCTCTGGAGTCC 843  
 QY 281 AsnAsp----- 282  
 DB 844 GGGACCCAGGAGTGTCTGGGCTGGGCTTCGAAAGCGGGTCCCGCGGGAGCGCTGGCGG 903  
 QY 282 ----- 282  
 DB 904 TTAGTAAAGAACCCGTCGGAACAATCAGAAAGTGGCTAGTCCCGGGCGGGGTGATTCG 963  
 QY 282 ----- 282  
 DB 964 ACTACGGGCGAGGGGGCGGGGTCAATTTAAAGTGGGTCCGGGATGGAGCTGCCCCAG 1023  
 QY 282 ----- 282  
 DB 1024 GGACTGACCTGGGCGGTGACTAGAAAGGTGAGACTAGATAGGGGGCGGGACGACACAA 1083  
 QY 282 ----- 282  
 DB 1084 GGCATAGTATGGATTGGATCAGTGGCATGACCCGACCGCCTGTCTTCAGTCCCCCTA 1143  
 QY 282 ----- 282  
 DB 1144 ACCTCGTCCCCCCCCCACAAGTGGAGTGGGAAGAGGAAGAGGAGGTGTGCT 1203  
 QY 282 ----- 282  
 DB 1204 GGGTCAAGGTCTGTAATCCCTGGCTTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1263  
 QY 282 ----- 282  
 DB 1264 TTTGCGACACAGGTTTCTCGTATAGCTCTGGCTGTCTGGAACTCACTTTGTAAGACA 1323  
 QY 282 ----- 282  
 DB 1324 GGCTGGCCCTCGAACTCAGAAATCCGCTCTCTGCTCTCTGCTCTCTGAGTGTGTGATTAAAGG 1383

QY 283 -----GluLeuPheGlnThrPheCys 289  
Db 1384 CGTGTGCCACTACCGCGCTACTGCCTGCTCTTCTTCTCCTCAGCTTTTCCAGACTTCTGT 1443  
QY 290 AspPheProGluLeuValAspValSerIleAsnGlnAlaProArgValGlyProAlaGly 309  
Db 1444 GACTTTCCGGAATCGTGAATGTCAGCATCAAGCAGC-CCACGTGTGTGGTCCGGCAGGG 1501  
QY 310 GluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhePro 329  
Db 1502 AG-CACCGCGTGTCTACTGTCCACGAGTGTGAGCGGCACATCTTGAAGCGGAGTTTCG 1560  
QY 330 GlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCys 349  
Db 1561 GGGCTCCCTCAGCGCTGTCTTTCGTGGCCCTCGTGGATGGGTACTTCCGCTGATCTGC 1620  
QY 350 AspSerArgHisTyrPheCysGlyGluValAlaProProArgLeuLeuGluGluAla 369  
Db 1621 GACTCCAGGCATTAATTCTGCAAGAGGTGTGGCGGCCACGGCTGTGGAGGAGGAGGCG 1680  
QY 370 AspValCysHisGlyProIleThr----- 377  
Db 1681 GAGCTGTGCCATGGACCCATCAGTTAGGAGCTAGTTGGGTCAACAGTGGGGATGGGGA 1740  
QY 378 -----LeuAspPhe 380  
Db 1741 TGGGCAGCGAGGACTGTCTTGTGGTGTGGACTGCTGCTGTTGGCTAGGTATAGACTTT 1800  
QY 381 AlaIleHisLeuIleGlyAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400  
Db 1801 GCATCCCAAGCTGAAGCGCGCTCGTCTCCAGGACCTATATTCTCCGCGCAGC 1860  
QY 401 ProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAsp 420  
Db 1861 CGCAGGACTATGACAGCTTCTCTTACCGCTGCTCCAGACTCTCTTGGCCCCGAC 1920  
QY 421 TyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440  
Db 1921 TACAAGGGGTGCTCATCCGCCAGGACCCAGCGGGCTTTCTCCCTGCTTGC-CTCAGC 1979  
QY 441 GlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSerGlyLeuArgVal 460  
Db 1980 CAGCCCCACAGAAAGCTCGCGAGGTGCTTGCAGC-TGCTGGAATTTC-GGGCTGCGAGTA 2037  
QY 461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGlyLysSerAsn 480  
Db 2038 GACGCTGTGCTGCTGAACCTAACATCTCTGCTCGCTCCAGACCCCAAGAAAGTCCCAAT 2097  
QY 481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500  
Db 2098 TTGATCGTGTGCGAAGGGCTGACACCCCGCGCTGCCCCCTGGCTGCTCCCGCTCTGC 2157  
QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu 520  
Db 2158 TGTGGCTGTGACACAGCTGAGCTTCCACAAATTCACACGGACAGCCTG---GGACACGAG 2214  
QY 521 AsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp 540  
Db 2215 AACCTGGTGTACGGTCTTTTACCAAGATCTTCCGTGGCCGAGCGGGAGTCTGTGAT 2274  
QY 541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560  
Db 2275 GTGAGACACATGACTCGGAAGTCTCTGGAAGTTCATGGACTCCAGACATCGGACTGC 2334  
QY 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580  
Db 2335 ATGGAGTCTTTTCTGGAAGCGCAGAGCTTGATGAGCCAAAGTATCTACCCGACCTGGT 2394  
QY 581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600  
Db 2395 TTACTGACGGCGTCTGCTGCTGGAGACAGCATCATGGTGCAGGAATTTGTGTATCTA 2454  
QY 601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeu 620

Db 2455 GGAGCAATTGACATGATACCTGCCAAGCGTGGCCACCTGGTGTACGACGCTGAAACTG 2514  
QY 621 GlnValThrLysGlnLeuAlaIleAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640  
Db 2515 CAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTACTTGGAGGACAAAGGCTTCTCTCAC 2574  
QY 641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnProPro 660  
Db 2575 GGCAACGTCTCAGCACGGAAGTGTCTCTGGCTCGTGGGGGGTGTATGGAAATCCACCT 2634  
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680  
Db 2635 TTCAITTAAGCTGAGTGTATCTGGTGTGCTGCCACCTGTGCTGAGCTGGAAATGCTCACC 2694  
QY 681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700  
Db 2695 GACAGAAATACCTGGTGGTGGCCCCGGAATGTCTCCAGAGGCTCAGACACTCTGCTTGGAG 2754  
QY 701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720  
Db 2755 GCTGACAACTGGGGCTTTTGGAGCCACCGTGGAGGTGTCTCAGCGGGGACCCGCGCCAC 2814  
QY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740  
Db 2815 ATCACTCGCTGGAGGCCGCCAAAGCTGAAGTCTTATGAGACCCAGGAGCAGCTGCC 2874  
QY 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
Db 2875 GCTCTCAATGACAGAACTGCGGGGACTTATCACACAGTGCATGGCGTATGATCTGTGC 2934  
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780  
Db 2935 CGGCGCCCTCTCTCCAGCTATCTCTCAGAGACCTCAACGGCTCATTTACATCAGATTAC 2994  
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800  
Db 2995 GAGCTCTCTCCA-GACCCCCACACTGGATCCCGAGTCCCGAGTCTCGAGATGAGCTGCGT- 3049  
QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyr 820  
Db 3050 GCGCGCCAGCTCTATGCTCTGCCAGACCCCGCCATATC-GAGGAGAGACACCTTAAGTAC 3108  
QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840  
Db 3109 ATCTCTTTCTGGGCAAGCGCAACTTTGGCAGCGTGGAGCTGTGCGCTATCACCCCTG 3168  
QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860  
Db 3169 ---GACAAATACGGGACCCCTGTGTGCGAGTGAACAGCTACAGCACAGCGGGCCAGACCAG 3225  
QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880  
Db 3226 CAGAGGGACTTCCAGCGGAGATTTCAGATCTTAAGGCTCTGCACAGGACTTCACTGCTC 3285  
QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900  
Db 3286 AAGTACCGGGAGTCAGCTATGGGCGCAGTCCAGAGCCCTGCGTTCGTGATGAGTAC 3345  
QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAsp 919  
Db 3346 CTGCCAGCGCTCGCTCGAGACTTCTGTGAGCGCATCTCTGAGCGCCATCGCGCGCCCTCACACCGAC 3405  
QY 920 ArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArg 939  
Db 3406 CGCTACTGCTGTTCTGCTTGGCAGATCTGCAGAGGCATGGAGTACCTGGGTGCGCGCCG 3465  
QY 940 CysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLys 959  
Db 3466 TGCCTACACCGCTGACCTGGCTGCGCGCAACATCTTGTGGAGAGCGAGGCTCATGTGAG 3525  
QY 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArg 979





815 Db :||||| 874  
260 QY GACCTGGAGCGCTGATCCAGCCGGGGCGCGGACCTTCCAGCTGGGCTCCCTGG 874  
875 Db :||| 277  
277 QY :||| 934  
935 Db :||| 994  
296 QY :||| 316  
995 Db :||| 1054  
316 QY :||| 336  
1055 Db :||| 1114  
336 QY :||| 356  
1115 Db :||| 1174  
356 QY :||| 376  
1175 Db :||| 1234  
376 QY :||| 1294  
1295 Db :||| 1354  
416 QY :||| 436  
1355 Db :||| 1414  
436 QY :||| 456  
1415 Db :||| 1474  
456 QY :||| 476  
1475 Db :||| 1534  
476 QY :||| 496  
1535 Db :||| 1594  
496 QY :||| 515  
1595 Db :||| 1654  
515 QY :||| 535  
1655 Db :||| 1714  
535 QY :||| 555  
1715 Db :||| 1774  
555 QY :||| 575  
1775 Db :||| 1834  
575 QY :||| 595  
1835 Db :||| 1894  
595 QY :||| 615

1895 Db GGAATTTGTACACCTGGGGCCATACATATATCTGCGAAACGTGGCCACCTGGTGCC 1954  
615 QY rlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyraLeuAAsnTrpLeuGluAs 635  
1955 Db AGCCAGCTGGAAGCTCAGGTGCTCAACAGCTGGCTACGCCCTCAACTATCTGGAGGA 2014  
635 QY pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655  
2015 Db CAAGAGCTTGGCCCATGGCAATGTCTCTCCGGAAGGTGCTCTGGCTGGGAGGGGGC 2074  
655 QY yAspGlyAsnProPheLeuLysLeuSerAspProGlyValSerProThrValLeuSe 675  
2075 Db TGATGGAGCCCGCCCTTCATCAAGCTGAGTGACCTTGGGTGAGCCCGCTGTGTAG 2134  
675 QY rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695  
2135 Db CTTGGAGATGCTCACCAGCAGGATCCCTGGGTGGCCCGCCGAGTGTCTCCGGAGGGCGCA 2194  
695 QY nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715  
2195 Db GACACTAGCTTGGAGAGCTGACAAGTGGGGCTTCGGCGCCACGGTCTGGGAAGTGTTAG 2254  
715 QY nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTrpGluAs 735  
2255 Db TGGCGTCAACATGCTGAGTGGCTGATCTCTGCTAAGAACTCCAAATTTATGAGGA 2314  
735 QY pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755  
2315 Db CCGGAGCAGAGCTGCGGGCCCGCCCAAGTGGACAGAGTGGCCCTGCTGATTCAACAGTGCAT 2374  
755 QY tAlaTyraAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775  
2375 Db GGCCTATGAGCGGTTCAGAGGCCCTCTCTCGAGCCGCTTCATTCGTGACCTCAATAGCCT 2434  
775 QY uIleThrSerAspTrpGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795  
2435 Db CATCTCTTCAGACTATGAGCTCTCTCAGACCCACACTGGTGGCCCTGGACCTCTGTGA 2494  
795 QY pGluLeuCysValAlaGlyAlaGlnLeuTyraAlaCysGlnAspProAlaIlePheGluGl 815  
2495 Db TGGGTGTGG--AATGGTGGCCAGCTCTATGCTGCTGCAAGACCCACAGCTCTTCGAGGA 2551  
815 QY uArgHisLeuLysTrpIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
2552 Db GAGACCTCAAGTACATCTCAGCTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTG 2611  
835 QY sArgTyraAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855  
2612 Db CCGCTATGACCCGCTAGCGACATAACAGGTGCCCTGGTGGCCCGTGAACAGCTGCAGCA 2671  
855 QY sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875  
2672 Db CAGCGGGCCAGACCAGCAGAGGAGCTTTCAGCGGAGATTTCAGATCCTCAAAGCACTGCA 2731  
875 QY sSerAspPheIleValLysTrpArgGlyValSerTyroGlyProGlyArgGlnSerLeuAr 895  
2732 Db CAGTGAATTCATTGTCAAGTATCGTGGTGTTCAGCTATGCGCCCGCCGCGAGCTGCG 2791  
895 QY gLeuValMetGluTyroLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGl 915  
2792 Db GCTGTCATGAGTACCTGCGCCAGCGGCTGTCTGGCGACTTCCTGACAGCGCGCACCGCGC 2851  
915 QY y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTy 934  
2852 Db CGCGCTCGATGCCAGCGCCCTCTCTCTATTCTTCGAGATCTGCAAGGCGCATGGAGTA 2911  
934 QY rLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954  
2912 Db CCTGGGTCCCGCGCTGCGTGCACCGGACCTGGCCGCCCGCAACATCTCTCGTGGAGAG 2971  
954 QY rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974  
2972 Db CGAGGCACAGTCAAGATCGCTGACTTCGGCTTAGCTGCTGCTGCTTGAACAAGA 3031

QY 974 pTyrValValArgGluProGlyGlnSerProIlePheThrPheTyrAlaProGluSerLe 994  
 Db 3032 CTACTACGGTGGTCCGAGCCAGCCAGCCAGCCATTTCTGTATGCCCCGAATCCCT 3091  
 QY 994 uSerAspAnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGI 1014  
 Db 3092 CTCGGACACATCTTCTCGCCAGTCAGACGCTCGAGCTCGGGTCTGCTGTACGA 3151  
 QY 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGI 1034  
 Db 3152 GCTCTTCACTACTCGACACAAAGCTCGAGCCCTCGCGAGTTCCTCGGATGATGG 3211  
 QY 1034 yProGluArgGluGlyProProLeuCysArgLeuLeuLeuAlaGluGlyArgAr 1054  
 Db 3212 ATGTGAGCGGATGTCCTCCCTCTGCGCCCTCTTGGAACTGCTGAGAGGCGCCAG 3271  
 QY 1054 gLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074  
 Db 3272 GCTGCGGCGCTCTGCTGCGCTGCTGAGGTG----- 3305  
 QY 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094  
 Db 3306 -----AGCGCCGAGGGCTAGCTTCAGTTCCAGTCTGTAGAT-----TG 3346  
 QY 1094 pArgGlyArgProGly 1099  
 Db 3347 GGCGGGGTCTCGGC 3362  
 RESULT 3  
 AAT30862  
 ID AAT30862 standard; DNA; 3807 BP.  
 XX AC AAT30862;  
 XX DT 13-SEP-1996 (first entry)  
 XX DE Protein tyrosine kinase JAK3 cDNA.  
 XX KW JAK3; protein tyrosine kinase; cell proliferation; differentiation;  
 KW signal transduction; leukaemia; aplastic anaemia; myelodysplasia;  
 KW polycythemia vera; thrombocytosis; gene therapy; diagnosis; ss.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 168..3359  
 FT /tag= a  
 FT polyA\_signal 3692..3696  
 FT /tag= b  
 FT polyA\_signal 3793..3797  
 FT /tag= c  
 XX PN W09618639-A1.  
 XX PD 20-JUN-1996.  
 XX PF 15-DEC-1995; 95WO-US16435.  
 XX PR 15-DEC-1994; 94US-0357598.  
 XX PA (UOJO ) UNIV JOHNS HOPKINS SCHOOL MED.  
 XX PI Civin CI, Safford MG, Small D;  
 XX DR WPI; 1996-300568/30.  
 XX P-PSDB; AAR96037.  
 XX PT Protein tyrosine kinase, JAK3, protein and nucleic acid - used in  
 PT the gene therapy of cellular proliferative diseases, e.g. leukaemia,  
 XX aplastic anaemia etc.  
 XX Claim 4; Page 43-47; 97pp; English.

XX A cDNA clone (AAT30862) codes for JAK3 (AAR96037), a new member of the  
 CC JAK family of non-receptor protein tyrosine kinases, that probably  
 CC plays a role in growth factor modulated differentiation,  
 CC proliferation and survival of haematopoietic stem/progenitor cells.  
 CC It was derived from CD34+ mRNA obtd. from normal human bone marrow  
 CC by PCR amplification using primers (see also AAT30863-64) based  
 CC on highly conserved motifs from protein tyrosine kinase catalytic  
 CC domains. The JAK3 coding sequence can be incorporated into a  
 CC vector and used for prodn. of recombinant JAK3. It can also be  
 CC used in gene therapy protocols for leukaemia, myelodysplasia,  
 CC polycythemia vera, thrombocytosis and aplastic anaemia, or to  
 CC stimulate haematopoietic cell proliferation. The gene was  
 CC localised to chromosome 19, band p12-13.1.  
 XX SQ Sequence 3807 BP; 716 A; 1219 C; 1099 G; 773 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 3807  
 Score: 4418.00 Matches: 870  
 Percent Similarity: 86.94% Conservative: 62  
 Best Local Similarity: 81.16% Mismatches: 130  
 Query Match: 75.39% Indels: 12  
 DB: 17 Gaps: 6  
 US-09-397-967A-16 (1-1099) x AAT30862 (1-3807)  
 QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
 Db 168 ATGGCACCTCCAAAGTGAAGAGAGCCCTCGATCCCTCAGGTTTCATGACGCTTCTGCC 227  
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
 Db 228 ACGGAGGCTGTGCCCTGCATGTGCTGCCCGCTCGGGCCCGGGCCCGCCAGCGC 287  
 QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLeuAla 60  
 Db 288 CTATCTTCTCTTGGGGACCACTTGTGCTGAGACCTGTGCGTGCAGGTGCGCAAGGCC 347  
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 Db 348 AGCGGATCTGCTGTGTACCTCCTCTTGTCTGGCCAGGAGGACCTCTCTCTGC 407  
 QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThr-GlnValLeuValTy 100  
 Db 408 TGGTT-CCCCCGAGCCACATCTTCTCGTGGAGGATGCCAGACCCCAAGTCTCTGTA 466  
 QY 100 rArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLe 120  
 Db 467 CAGGATTGCTTTTACTTCCCAATTCGTTGGGCTGGAGAAGTGCACCGCTTCGGGCT 526  
 QY 120 uArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGl 140  
 Db 527 ACGCAAGATTGGCCAGTGTCTATCTTGACCTGCCAGTCTGAGGACCTCTTTGCCCA 586  
 QY 140 nHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGl 160  
 Db 587 GCACCGCAGTGACCTGTGTAGTGGGCGCTCTCCCGTGGCTCAGTCTCAAGGACGAGG 646  
 QY 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluAlaGlnAr 180  
 Db 647 TGAGTGTCTCAGCTGCGCTGTTGGACCTCGCCCGGATGGCGAGAGAGGCCAGCG 706  
 QY 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200  
 Db 707 GCGGGGAGAGTGTGTAGACTGTTCAGCTACAGGCTGCTACCCCAAGCTCTGCGGA 766  
 QY 200 pValIleGlnGlnAenPheValThrArgArgIleArgThrValValLeuAl 220  
 Db 767 CTTGATCCAGGCGCTGAGCTTCGTGACGGGGAGGCGTATTTCGGAGGACGGTGAGAGGCC 826  
 QY 220 aLeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleL 240



955 luAlaHisVallyllylleAlaAappPheGlyLeuAlaLysLeuProLeuGlyLysAspT 975  
3037 AGGCACACGTCAAGATCGCTGACTTCGGCTAGCTAAGCTGCTGCGCTTGACAAAGACT 3096  
975 yTyrrValValArgGluProGlyGlnSerProillePheTrpTyrAlaProGluSerLeuS 995  
3097 ACTACGTGGTCCGAGCAGAGCCAGAGCCCAATTTCTGGTATGCCCGCATCCCTCT 3156  
995 exAspAsnillePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluL 1015  
3157 CGGACAACTCTTCTCTCGCCAGTCAGAGCTCTGGAGCTTCGGGGTCTGCTCTACGAGC 3216  
1015 euPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuAtrGmetMetGlyP 1035  
3217 TCTTCACTACTCGCAAAAGTGCAGCCCTCGCCCGAGTTCCTCGCGATGATGGAT 3276  
1035 roGluuArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgL 1055  
3277 GTGAGCGGGATGCCCGCTCTGCGCCCTCTTGGAACTGCTGGAGAGGGCCAGAGGC 3336  
1055 euProProProProThrCysProThrGluVal 1065  
3337 TGC CGCGGCTCTTGTGCTGCTGCTGAGGTG 3368  
RESULT 4  
AAV61801  
ID AAV61801 standard; cDNA; 3435 BP.  
XX AAV61801;  
AC  
XX 20-JAN-1999 (first entry)  
DT  
XX JAK2 protein encoding cDNA.  
DE  
XX JAK2 protein; cytoplasmic domain; betac subunit; screening; asthma;  
KW interleukin; granulocyte macrophage-colony stimulating factor; GM-CSF;  
KW IL-3; IL-5; human; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 10..3429  
XX /\*tag= a  
XX /product= "JAK2 protein"  
XX  
XX WO9843087-A1.  
XX  
XX 01-OCT-1998.  
XX  
XX 23-MAR-1998; 98WO-US05387.  
XX  
XX 24-MAR-1997; 97US-0041511.  
XX  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX  
XX Kaytes PS, Roberds SL;  
XX  
XX WPI; 1998-532151/45.  
XX P-PSDB; AAW76425.  
XX  
XX Screening for compounds useful for preventing or treating asthma -  
XX by determining if compounds inhibit binding of the JAK2 protein to  
XX e.g. IL-3, IL-5 or GM-CSF  
XX  
XX Disclosure; Fig 22A-C; 112pp; English.  
XX  
XX This cDNA encodes the full-length JAK2 protein. The N-terminal fragment  
XX of JAK2 protein can be used in the method of the invention of screening  
XX for compounds useful for treating or preventing asthma. The method  
XX comprises contacting a molecule comprising at least the N-terminal 294  
XX amino acid residues of the JAK2 protein, with another molecule comprising  
XX at least 13 membrane-proximal cytoplasmic amino acids of interleukin  
XX (IL)-3, IL-5 or granulocyte macrophage-colony stimulating factor (GM-CSF)  
XX

CC proteins in the presence of the candidate compound, and determining  
CC whether the first and the second molecules form a complex. If the  
CC compound inhibits complex formation, it can be used to treat asthma.  
XX  
SQ Sequence 3435 BP; 1157 A; 598 C; 733 G; 947 T; 0 other;  
Alignment Scores: 3.07e-191 Length: 3435  
Pred. No.: 2629.00 Matches: 534  
Score: 66.21% Conservative: 193  
Percent Similarity: 48.63% Mismatches: 340  
Best Local Similarity: 44.86% Indels: 31  
Query Match: 19 Gaps: 14  
DB: 14  
US-09-397-967a-16 (1-1099) x AAV61801 (1-3435)  
QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProGlu 39  
Db 111 TCCAGTTCTTTCAGT-----GTATCTTTTACCATTCCCTTGGGMAATCTGAGGCAGA 161  
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaLal 59  
Db 162 TTAATCTGACCTTCCATCTGGGGAGTATGTCAGAGAAATCTGTATTGCTGCTCTTAA 221  
QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79  
Db 222 AGCTTGTGGTATCACACCTGTATCATATAATATGTTTGTCTTATGAGTGAACAGAAAG 281  
QY 79 rCysTrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99  
Db 282 GATCTGGTATCCACCCCAACCATGCTTCCATATAGATGATGATCAACAGGACATATGTACT 341  
QY 99 lTyrArgLeuArgPheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisAr 117  
Db 342 CTACAGAAATAGATTTTACTTCTCTGTTGATGCGAGTGGCAGCAACAGAGCTATCG 401  
QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137  
Db 402 GCATGAATATCTCGAGGTGCTGAAGCTCTCTTCTTGTATGACTTTGTCTATGCTTCTACT 461  
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157  
Db 462 CTTTGCTCAGTGGCGGATGATTTTGTGATGGATGATATAAAAGTCTGTGACTCTGA 521  
QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177  
Db 522 AACACAGGAAGATGTCTTGGATGGCAGTCTTAGATATGATGAGAAATAGCAAGAAAA 581  
QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe 197  
Db 582 CGATCAAAACCCCACTGGCCATCTATACTTATCAGCTACAAGACATTTTACCAGAAATG 641  
QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgGlyLeuArgThrVa 217  
Db 642 TATTGAGCAAGATCCAAAGACTATCATTTTTCAGAGGAGCAAGATTAAGTACAGATT 701  
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237  
Db 702 TCGCAGATTATTTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAAT 761  
QY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl 257  
Db 762 GTATCTTAAATCTGGAACCTCTGAGTCTGCTTCTACACAGAGAAATTTGAAAGTAAA 821  
QY 257 yLeuProGlyAla-----GlnGluLeuProGlyLeuLeuArgValAlaGlyAs 273  
Db 822 AGAACCTGGAAGTGTCTCTTCAAGTGGAGAGATTTTTCAGCAATTCAGCAATTCAGCAAT 881  
QY 273 pAsnGlyIleProTyrTrpSer-----AsnAspGluLeu 284  
Db 882 CGGTGGAATTCAGTGTTCAGAGGGGAAACATAAGAAAGTGAAGACACTGACAGAACAGGA 941  
QY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304



QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuAraMetMetGlyProGluar 1037  
 Db 3150 ATACATTGAGAGAGTAAAGTCCACAGCGGAATTTATCGGTATGTCGCAATGACAA 3209  
 QY 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPr 1056  
 Db 3210 ACAAGGACAGATGATCGTGTTCATTTGATAGAACCTTTGAGAAATATGGAAGATTACC 3269  
 QY 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuLeuCysTrpAlaProGl 1076  
 Db 3270 AGACACAGATGGATGCCAGATGAGATCTATATGATCATGACAGAAATGCTGGAACAATAA 3329  
 QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093  
 Db 3330 TGTAAATCAACGCCCTCTTTAGGATCTAGCTTTTCGAGTGGATCAATA 3381

RESULT 5  
 ACC46252  
 ID ACC46252 standard; cDNA; 6037 BP.  
 XX  
 AC ACC46252;  
 XX  
 DT 02-JUN-2003 (first entry)  
 XX  
 DE Human dithp intracellular signalling protein-encoding cDNA.  
 XX  
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;  
 KW cancer; cell proliferative disorder; autoimmune disorder;  
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;  
 KW neurological disorder; gastrointestinal disorder; transport disorder;  
 KW connective tissue disorder; drug screening; proteome analysis;  
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;  
 KW disease model; toxicological testing; transcript imaging;  
 KW intracellular signalling; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200297031-A2.  
 XX  
 XX 05-DEC-2002.  
 XX  
 XX 27-MAR-2002; 2002WO-US10056.  
 XX  
 XX 28-MAR-2001; 2001US-279619P.  
 XX 29-MAR-2001; 2001US-280067P.  
 XX 29-MAR-2001; 2001US-280068P.  
 XX 16-MAY-2001; 2001US-291280P.  
 XX 17-MAY-2001; 2001US-291829P.  
 XX 17-MAY-2001; 2001US-291849P.  
 XX 19-JUN-2001; 2001US-299428P.  
 XX 20-JUN-2001; 2001US-299776P.  
 XX 20-JUN-2001; 2001US-300001P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;  
 XX Daughtery SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin BH;  
 XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 XX Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;  
 XX  
 XX WPI; 2003-129518/12.  
 XX P-PSDB; ABR41310.  
 XX  
 XX Novel human diagnostic and therapeutic polypeptide useful for  
 XX identifying test compound which specifically binds to a polypeptide  
 XX encoded by human diagnostic and therapeutic polynucleotide, and to  
 XX induce antibodies  
 XX  
 XX Claim 2; SEQ ID No 173; 591pp; English.  
 XX  
 XX The invention relates to novel human diagnostic and therapeutic

CC polynucleotides designated dithp (ACC46080-ACC46749) and to their  
 CC encoded proteins (DITHP; ABR41136-ABR41812). The invention also relates  
 CC to polynucleotide sequences at least 90% identical to the dithp cDNA  
 CC sequences of the invention; recombinant vectors, host cells and  
 CC transgenic organisms comprising a dithp nucleic acid sequence; the  
 CC recombinant production of DITHP proteins; antibodies specific for DITHP  
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods  
 CC of detecting dithp nucleotide and protein sequences; methods of screening  
 CC for compounds which specifically bind a DITHP protein; and methods of  
 CC assessing the toxicity of test compounds using a dithp hybridisation  
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the  
 CC diagnosis of a wide variety of conditions including cancer and other cell  
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,  
 CC viral, fungal or parasitic infections; gastrointestinal disorders; metabolic  
 CC disorders; neurological disorders; gastrointestinal disorders; transport  
 CC disorders; and connective tissue disorders. They may also be used to  
 CC screen for modulators of protein activity or gene expression. DITHP  
 CC proteins can additionally be used in analysis of the proteome of a tissue  
 CC or cell type and to induce antibodies. The dithp nucleic acids are  
 CC additionally useful in somatic or germline gene therapy of the disorders  
 CC mentioned above, as a source of antisense sequences, as a source of  
 CC probes and primers, in genotyping and identification of individuals, in  
 CC the generation of transgenic animal models of human disease or knock in  
 CC humanised animals, in toxicological testing, and in transcript imaging.  
 CC The present sequence represents a dithp cDNA encoding a DITHP protein  
 CC which has intracellular signalling activity.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

Sequence 6037 BP; 1959 A; 1011 C; 1206 G; 1861 T; 0 other;

Alignment Scores: 9.33e-191 Length: 6037  
 Pred. No.: 2627.00 Matches: 534  
 Score: 66.12% Conservative: 192  
 Percent Similarity: 48.63% Mismatches: 341  
 Best Local Similarity: 44.83% Indels: 31  
 Query Match: 25 Gaps: 14  
 DB:

US-09-397-967A-16 (1-1099) x ACC46252 (1-6037)

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProGlu 39  
 Db 512 TCCAGTTCTTCAGGT-----GTATCTTACCATTCCTCGGAAATCTGAGGCAGA 562  
 QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59  
 Db 563 TTATCTGACCTTTCCATCTGGGGAGTATGTTGCAGAAAGAAATCTGTATTGCTGCTTCTPA 622  
 QY 59 sAlaCysGlyTleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79  
 Db 623 AGCTTGTGGTATCACCTGCTGTATCATATATGTTGCTTTAATGAGTGAACAGAGAAG 682  
 QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99  
 Db 683 GATCTGGTATCCACCCACCATGCTTCCATATAGATGATGATCAACAGGCATATGTACT 742  
 QY 99 lTyArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117  
 Db 743 CTACAGATAAGATTTTACTTTCTCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802  
 QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137  
 Db 803 GCATGGAATATCTCGAGGTGCTGAGCTCTCTTCTTCTGATGATTTGTCTGCTTCTTACT 862  
 QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157  
 Db 863 CTTTGCTCAGTGGCGGCATGATTTTGTGTCACGGATGATGATGATGATGATGATGATGATG 922  
 QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177  
 Db 923 AACACAGGAAGATGCTTTGGGATGGCAGTGTATGATGATGATGATGATGATGATGATGATG 982





3131 CATTTGTAAGTACAGGAGTGTCTACAGTGTCTGCGGTAATCTAAATTAATAT 3190  
 898 tGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHi 917  
 3191 GGAATATTTACCATATGGAAGTTTACGAGACTATCTTCAAAAACATAAAGAACGGATAGA 3250  
 917 sThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAl 937  
 3251 TCACATAAAACTTCTCGAGTACACATCTCAGATATGCAAGGGATGGAGTATCTTGGTAC 3310  
 937 aArgArgCysValHisArgAspLeuAlaIaArgAsnIleLeuValGluSerGluAlaHi 957  
 3311 AAAAAGGTATATCCACAGGATCTGGCAGCAGAAATATATTTGGTGAGAACGAGAACAG 3370  
 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTrpVa 977  
 3371 AGTGAATAATGGAGATTTTGGGTTTAAACCAAGTCTTGCCACAGACAAAGAACTACTATA 3430  
 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997  
 3431 AGTAAAGAACCTTGGTGAAGTCCCATTTCTGGTATGCTCCAGATCATCTGACAGAGAG 3490  
 997 nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017  
 3491 CAAGTTTCTGGCCTCAGATGTTTGGAGCTTGGAGTGTCTCTGATGAACCTTTTCAC 3550  
 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr 1037  
 3551 ATACATTGAGAAGAGTAAAGTCCACCCAGCGGAAATTTATGCTATGATGGCAATGACAA 3610  
 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr 1056  
 3611 ACAAGGACAGATGATGTTTCCATTTGATAGAACTTTTGAAGATATATGGAAGATTACC 3670  
 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076  
 3671 AAGACACAGATGATGCCAGATGAGATCTATATGATCATGACAGAAATGCTGGAACAATA 3730  
 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093  
 3731 TGTAATCAACGCCCTCTTTTAGGGATCTAGCTCTTTCGAGTGGATCAAAATA 3782  
 RESULT 6  
 AAX80971  
 ID AAX80971 standard; cDNA; 4482 BP.  
 XX  
 AC AAX80971;  
 XX  
 DT 03-SEP-1999 (first entry)  
 XX  
 DE Human JAK2 kinase encoding cDNA.  
 XX  
 KW JAK2 kinase; arteriosclerosis; asthma; bronchitis; emphysema; psoriasis;  
 KW inflammatory bowel disease; inflammation; osteoarthritis; oncogenesis;  
 KW rheumatoid arthritis; septic shock; systemic lupus erythematosus;  
 KW leukaemia; human; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN US5914393-A.  
 XX  
 PD 22-JUN-1999.  
 XX  
 PF 05-DEC-1995; 95US-0567508.  
 XX  
 PR 05-DEC-1995; 95US-0567508.  
 XX  
 PA (INCY-) INCYTE PHARM INC.  
 XX  
 PI Coleman R, Stuart SG;  
 XX  
 WI 1999-384188/32.  
 DR

DR P-PSDB; AAV21698.  
 XX Polyptides and amino acids useful for modulating human jak2 kinase  
 PT activity  
 XX  
 XX Claim 2; Columns 23-28; 37pp; English.  
 CC This cDNA encodes a human JAK2 kinase polypeptide. Host cells transformed  
 CC with recombinant jak2 kinase nucleic acid are used for the recombinant  
 CC production of the protein. Purified JAK2 may be used to produce  
 CC antibodies or identify antagonists or inhibitors of JAK2. JAK2, anti-JAK2  
 CC antibodies and JAK2 antagonists or inhibitors may be used to treat,  
 CC prevent or diagnose conditions associated with altered or uncontrolled  
 CC jak2 expression, e.g. arteriosclerosis, asthma, bronchitis, emphysema,  
 CC inflammatory bowel disease, inflammation, leukemia, oncogenesis,  
 CC osteoarthritis, psoriasis, rheumatoid arthritis, septic shock and  
 CC systemic lupus erythematosus.  
 XX  
 XX Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1.8e-190 Length: 4482  
 Score: 2621.00 Matches: 532  
 Percent Similarity: 66.03% Conservative: 193  
 Best Local Similarity: 48.45% Mismatches: 342  
 Query Match: 44.73% Indels: 31  
 DB: 20 Gaps: 14  
 US-09-397-967a-16 (1-1099) x AAX80971 (1-4482)  
 QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39  
 Db 447 TCCAGTTCTTCAGGT-----GTATCTTTTACCATTCCCTTGGGAAATCTTGAGGCAGA 497  
 QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59  
 Db 498 TTAATCTACCTTTCCATCTGGGGAGTATGTTGGGAAAGAAATCTGTAATGCTGCTCTCAA 557  
 QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79  
 Db 558 AGCTTGTGATATCACACTGTGATCATATAATATGCTTTTAAATGAGTGAACAGAAAG 617  
 QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99  
 Db 618 GATCTGCTATCCACCAACCATGCTTCCATATAGATGAGTCAACACAGGCATATGTACT 677  
 QY 99 lTyrArgLeuArgPheTyrPhePheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117  
 Db 678 CTACAGAAATAGATTTTACTTTCTCTGTTGATGATGATGATGATGATGATGATGATGATG 737  
 QY 117 gpheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137  
 Db 738 GCATGGAATATCTCGAGGTGCTGAAGCTCTCTCTTCTTGTGATGATGATGATGATGATGAT 797  
 QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157  
 Db 798 CTTTGTCTCAGTGGCGCATGATTTTGTGTCATGGATGGATAAAAGTACCTGTGATCATGA 857  
 QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177  
 Db 858 AACACAGGAAGAATGCTTGGGATGACAGTGTGATGATGATGATGATGATGATGATGATGATG 917  
 QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe 197  
 Db 918 CGATCAAAACCCACTGGCCATCTATACTCTATCAGCTACAGATACAAAGACATTTTACCACA 977  
 QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgGlyIleArgArgThrVa 217  
 Db 978 TATTGCGAAAGATCCCAAGACTATCATATTTTTCAGAGGAGCGGAAATAGGTACAGATT 1037  
 QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237  
 Db 1038 TCGCAGATTATTATCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAAT 1097



QY	237	sTYrIlleuAspIeuGluArgLeuHieProAlaAlaThrThrGluThrPheArgValGl	257
DB	1098	GTATCTTATAAATCTGGAACACTCTGCAGTCTGCCTTCTACACAGAGAAATTTGAAGTAAA	1157
QY	257	YLeuProGlyAla-----GInGluLuproGlyLeuLeuArgValAlaGlyAs	273
DB	1158	AGAACCTGGAAGTGGTCTCTCAGGTGAGAGATTTTGCAACCATTTATACTACTGGAAA	1217
QY	273	pAsnGlyIleProTrpSerSer-----AsnAspGluLe	284
DB	1218	CGGTGGAAATTCAGTGGTCAAGAGGGAACATAAAGAAAGTCAGACACTGACAGAACACGGA	1277
QY	284	uPheGlnThrPheCysAspPheProGluIleValAspValSerIleGlnAlaProAr	304
DB	1278	TTTACAGTTATATTGGCAATTTCTTAATTATTGATGTCTAGTATTAAAGCAAGCAACCA	1337
QY	304	gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe	324
DB	1338	AGAGGGT--TCAAATGAAGCCGAGTGTAACTATCCATAAGCAAGATGTTAAANATCT	1394
QY	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy	344
DB	1395	GGAAATTTGAATCTAGCTCAATTAAGGAAGCTTTGTCTTTCTGTGCTAATTAATTGATGATA	1454
QY	344	rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe	364
DB	1455	TTATAGATTAACTGCAGATGCATCATCATTAACCTCTGTAAGAAGATGCACCTCCAGCCGT	1514
QY	364	uLeuGlnGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy	384
DB	1515	GCCTGAAAATATACAAAGCAACTGTCATGCCCAATTCGATGGATTTTGCCATTAGTAA	1574
QY	384	sLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTy	404
DB	1575	ACTGAAGAAAGCAGGTAATCAGACTGCATGTATGTACTTCGATGCAGTCTCAAGGACTT	1634
QY	404	rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIysGlyCy	424
DB	1635	TAATAATATTTTGTCTTTGCTTGTGAGCGAGAAAATGTCATTGAATATAAACACTG	1694
QY	424	sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr	444
DB	1695	TTTGATTACAAAAATGAGAATGAAGAGTACAACTCAGTGGCAACAAGAGAACTTCAG	1754
QY	444	gSerLeuArgGluLeuLeuAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl	464
DB	1755	CAGTCTTAAAGATCTTTTGAATTTGTACCAGATGGAAACTGTTTCGCTCAGACAATAAT	1814
QY	464	aLeuTyLeuLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleVal	484
DB	1815	TTTCCAGTTTACTAAATGCTGTCCCCCAAGCCAAAGATTAATCAAACTCTTAGTCTT	1874
QY	484	lArg--ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaLe	503
DB	1875	CAGAAGCAATGGTGTCTCTGATGTACCACCTCCACCAACATTACAGAGGCTACTCATAT	1934
QY	503	uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGl	523
DB	1935	GAACCAAAATGGTGTTCACAAAATCAGAAATGAAGATTTTGATATTTAATGAAGCCCTTG	1994
QY	523	yHisGlySerPheThrIysIlePheArgGlySerArgGluValValAsp---GlyGl	542
DB	1995	CCAAGGCATTTTACAAAGATTTTAAAGCGCTCAGNAGAAAGTAGGAGACTTACGGTCA	2054
QY	542	uThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGl	562
DB	2055	ACTGCATGAACAGAAAGTCTCTTTTAAAGTTCTGGATAAAGCACACAGGACTATTTCAGA	2114
QY	562	uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLe	582
DB	2115	GTCTTCTTTGAAGCAGCAAGTATGATGACCAAGCTTCTCAAGCATTTGGTTTAAA	2174

QY	582	uHisGlyValCysMetalGlyAsp---SerIleMetValGlnGluPheValTyrLeuG1	601
Db	2175	TTATGGAGTATGTGTTCTGTGGAGACGAGAAATATTTCTGTTTCAGCAGTTTGTAAAAATTTGG	2234
QY	601	yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerXlaSerTrpLysLeuG1	621
Db	2235	ATCCTACGATACATATCTGAAAAAGAAATAAAATTTGTATAAATATATATTGGAACACTTGA	2294
QY	621	nValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisG1	641
Db	2295	AGTTGCTAAACAGTTGGATCGGGCCATGTCATTTCTAGAGAAACACCCTTATTCATGG	2354
QY	641	yAsnValSerXlaaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnPr	659
Db	2355	GAATGTATGTGCCAAAAATATCTGCTTATCAGAGAAAGACAGAGGAGACAGGAAATCC	2414
QY	659	oPropHeIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe	679
Db	2415	TCCTTTTCATCAACTTAGTATCTCGCATTTAGTATTACAGTTTGTGCCAAAGGACATCTCT	2474
QY	679	uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe	699
Db	2475	TCAGGAGAGATACCATGGGTACCACTGAATGCATTTGAAATTCCTTAATAATTTAAATTT	2534
QY	699	uGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl	719
Db	2535	GGCAACACACAAATCGAGTTTGTGTACCACCTTTGTGGAAATCTGCAGTGGAGGAGATAA	2594
QY	719	aHisIleThrSerLeuGlnProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLe	739
Db	2595	ACCTCTAAGTGCTCTGGAATCTCAAGAAAGACTACAAATTTTATGAAGATAGGCATCAGCT	2654
QY	739	uProAlaLeuLysTrpThrGluLeuAlaGlyLeuLeThrGlnCysMetalTyrAspPr	759
Db	2655	TCCTGCCAAAGTGGCAGAAATAGCAACACTTATAAATAATTTGTATGATTTATGAACC	2714
QY	759	oGlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs	779
Db	2715	AGATTTTCAGGCCTTCTTTTCAGAGCATCATCAGAGATCTTTAACAGTTTGTATTACTCCAG	2774
QY	779	pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy	798
Db	2775	TTATGAACTATTACAGAA---AATGACATGTTTACCAAAATATCAGGATAGTGCCTTGGG	2831
QY	798	sValAlaGlyAlaGlnLeuTyrXlaCysGlnAspProAlaIlePheGluClnArgHisLe	818
Db	2832	GTTTTCGTGTGCC-----TTTGAGACCCGGGATCTCTACACGTTTTGAAGAGAGACATTT	2885
QY	818	uLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAs	838
Db	2886	GAATTTTCTACAGCAACTTGGCAAGGTTAATTTTGGAGTGTGGAGATGTCCCGGTATCGA	2945
QY	838	pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr	858
Db	2946	CCCTCTACAGCAACAACCTGGGAGGTGGTGCCTGTAAAAAAGCTTCAGCATAGTACTGA	3005
QY	858	oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLysAlaLeuHisSerAspPh	878
Db	3006	AGAGCCTTAAGAGACTTTGAAAGGGAATTTGAATCTCTGAATCCCTCAGCATGACAA	3065
QY	878	eIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMe	898
Db	3066	CATTGTAAGTACAAGGAGTGTCTACAGTGTGTCGCGCTAATCTTAAAAATAAAGACGATAGA	3125
QY	898	tGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHi	917
Db	3126	GGAAATATTTACCATATGGAAGTTTACGAGACTATCTTCAAAAACATAAAGACGATAGA	3185
QY	917	sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAl	937
Db	3186	TCATATAAACTTCTCGATGACACATCTCAGATATGCAAGGTATGGAGATATCTTTGGTAC	3245
QY	937	aArgGlyCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHi	957

```

DB 3246 AAAAGGTATATCCAGGGATCTGGCAACGAGAAATATATTTGGTGAGAACGAGACAG 3305
QY 957 sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVa 977
DB 3306 AGTTAAATTTGGRGATTTTGGGTTTAAACCAAGTCTTGGCCACAGACAAAGAACTACTATAA 3365
QY 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997
DB 3366 AGTAAAGAACCTGGTGAAGTCCCATTTCTGGTATGCTCCAGAAATCACTGACAGAGAG 3425
QY 997 nIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh 1017
DB 3426 CAAGTTTCTGTGSCCTCAGATGTTTGGAGCTTTGGAGTGTTCTGTATGAACCTTTTCAC 3485
QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr 1037
DB 3486 ATACATTGAGAAGAGTAAAGTCCACAGCGGAATTTATCGTATGATTGGCAATGACAA 3545
QY 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr 1056
DB 3546 ACAGGGACAGATGATCGTGTTCATTTGATAGAACTTTTGAGAAATATGGAAGATTACC 3605
QY 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076
DB 3606 AAGACCAGATGGATGCCAGATGAGATCTATATGATCATGACAGAAATGCTGGAAACAATA 3665
QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093
DB 3666 TGTAATCAAGCCCTCTTTAGGGATCTAGCTCTTCGAGTGGATCAATA 3717

```

## RESULT 7

AAZ58947

ID AAZ58947 standard; cDNA; 4482 BP.

XX AC AAZ58947;

XX XX 03-MAY-2000 (first entry)

XX DE Human JAK2 kinase (HJAK2) encoding cDNA.

XX KW Janus family nonreceptor protein-tyrosine kinase-2; Jak2; HJAK2; human;  
 KW signal transduction; arteriosclerosis; asthma; bronchitis; emphysema;  
 KW inflammatory bowel disease; leukemia; oncogenesis; osteoarthritis;  
 KW psoriasis; rheumatoid arthritis; systemic lupus erythematosus;  
 KW cytosolic; osteopathic; dermatological; antibacterial; septic shock;  
 KW immunosuppressive; ss.

XX OS Homo sapiens.

XX XX US6019966-A.

XX XX 01-FEB-2000.

XX XX 19-NOV-1998; 98US-0196480.

XX XX 05-DEC-1995; 95US-0567508.

XX XX (INCY-) INCYTE PHARM INC.

XX XX Stuart SG, Coleman R;

XX XX WPI; 2000-146859/13.

XX XX P-PSDB; AAY77552.

XX PT Human Janus family nonreceptor protein-tyrosine kinases useful as  
 PT diagnostic reagents and for preventing, diagnosing and treating  
 PT diseases such as arteriosclerosis, asthma and leukemia -

XX PS Example 5; Fig 1A-F; 33pp; English.

XX CC This cDNA encodes a human Janus family nonreceptor protein-tyrosine  
 CC kinase-2 (Jak2) polypeptide (HJAK2). The Jak2 polypeptides may be used

CC as diagnostic reagents as they react with a range of target proteins  
 CC including growth hormone, prolactin, erythropoietin and cytokine  
 CC receptors. They may also be used for the production of antibodies  
 CC specific for Jak2, which may be used to inhibit its activity and prevent  
 CC or treat disorders associated with over expression of Jak2. Conversely,  
 CC the Jak2 polypeptide may be administered to supplement the patients own  
 CC production and counter mutations that may lead to the expression of an  
 CC inactive enzyme. The protein may also be used to screen candidate  
 CC reagents for modulators of Jak2 function. The antagonists and antibodies  
 CC bind to the Jak2 protein and prevent the transfer of high energy  
 CC phosphate molecules, therefore blocking signal transduction. Disorders  
 CC that may be treated by administration of Jak2 polypeptides, anti-Jak2  
 CC antibodies and the agonists and/or antagonists, include arteriosclerosis,  
 CC asthma, bronchitis, emphysema, inflammatory bowel disease, leukemia,  
 CC oncogenesis, osteoarthritis, psoriasis, rheumatoid arthritis, septic  
 CC shock and systemic lupus erythematosus.

XX SQ Sequence 4482 BP; 1477 A; 795 C; 953 G; 1249 T; 8 other;

## Alignment Scores:

Pred. No.: 1.8e-190 Length: 4482  
 Score: 2621.00 Matches: 532  
 Percent Similarity: 66.03% Conservative: 193  
 Best Local Similarity: 48.45% Mismatches: 342  
 Query Match: 44.73% Indels: 31  
 DB: 21 Gaps: 14

US-09-397-967A-16 (1-1099) x AAZ58947 (1-4482)

```

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39
DB 447 TCCAGTTCCTCAGGT-----GTATCTTTACCATTCCTCCCTGGGAAATCTGAGGAGA 497
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59
DB 498 TTATCTGACCTTTCCATCTCGGGAGTATGTTGGAGAGAAATCTGTATGCTGCTCTTAA 557
QY 59 sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
DB 558 AGCTTGTGGTATCACACCTGTATCATATAATATCTTTGCTTTAATGATGAAACAGAAAG 617
QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
DB 618 GATCTGTATTCACCCCAACCATGCTTCATATAGATGAGTCAACAGGCAATATGTAAT 677
QY 99 lTyrArgLeuArgPheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisAr 117
DB 678 CTACAGATATAGATTTTACCTTCTCGTGTATGTCAGTGGCAGCAACAGAGCCTATCG 737
QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
DB 738 GCATGGAATATCTCGAGGTGCTGAAGCTCTCTTCTGATGACTTTGTCTATGCTTACT 797
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
DB 798 CTTTGTCTCAGTGGCGGCATGATTTTGTGCATGGATGGAATAAAGACTCTGTGACTCATGA 857
QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
DB 858 AACACAGGAAGAATGTCTTGGATGACAGATGATAGATATGATGAGAAATAGCAAGAAA 917
QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe 197
DB 918 CGATCAAAACCCACTGGCCATCTATACTCTATCAGCTACAGACATCTTTACCACAATG 977
QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgGlyLeuArgThrVa 217
DB 978 TATTGACCAAGATCCCAAGACTATCATATTTTGACAAGAGAGCGAAATAGTACAGATT 1037
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237
DB 1038 TCGCAGATTATTATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCA 1097

```

Qy	237	s	Tyrlleu	leu	asp	leu	glu	arg	leu	his	pro	ala	ala	thr	thr	glu	thr	phe	arg	val	gl	257			
Db	1098	g	t	a	t	c	t	t	a	a	a	a	a	c	t	c	t	c	a	c	a	1157			
Qy	257	y	leu	pro	gl	y	ala	-----	g	l	n	g	l	u	p	r	o	g	l	y	leu	arg	273		
Db	1158	a	g	a	a	c	t	g	a	g	a	g	a	g	a	g	a	g	a	g	a	g	1217		
Qy	273	p	as	n	g	l	y	i	l	e	p	r	o	t	r	p	s	e	r	-----	a	s	n	284	
Db	1218	c	g	g	t	g	a	a	t	c	a	g	t	g	g	t	c	a	a	g	a	a	1277		
Qy	284	u	p	h	e	l	n	t	r	p	h	e	c	ys	a	s	p	p	h	e	pro	g	l	u	304
Db	1278	t	t	t	a	c	a	g	t	t	a	t	t	t	c	c	a	a	t	t	t	t	t	1337	
Qy	304	g	val	g	l	y	pro	ala	g	l	y	his	arg	leu	val	thr	val	thr	arg	met	asp	g	l	324	
Db	1338	a	g	a	g	g	t	---	t	c	a	a	t	g	a	a	c	c	g	a	g	i	t	1394	
Qy	324	u	g	l	u	ala	g	l	u	p	h	e	pro	g	l	y	leu	pro	g	l	u	ala	leu	344	
Db	1395	g	a	a	a	t	t	g	a	a	c	t	c	a	t	t	a	a	g	a	a	c	t	1454	
Qy	344	r	p	h	e	arg	leu	leu	c	ys	a	s	p	ser	arg	his	tyr	p	h	e	c	ys	364		
Db	1455	t	t	a	t	a	g	a	t	t	a	a	c	a	t	c	a	t	a	t	t	a	c	1514	

[illegible]

3246	AAAAAGGTATATCCACAGGGATCTGGCAACGAGAAATATATTGGTGGAGAACGAGAACAG	3305
Qy	svallvsiieAlaaspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrVa	977
Db	AGTTAAATTTGGRGATTTTCGGTTAAACCAAGTCTTGCCCAACAGACAAAGAACTATATAA	3365
Qy	lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs	997
Db	AGTAAAGAACCTGGTGAAGTCCCATTTCTGGTATGCTCCAGATCATCTGACAGAGAG	3425
Qy	nlIlePheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheTh	1017
Db	CAAGTTTCTGGCCTCAGATGTTTGGAGCTTTGGAGTGGTCTGTATGAACTTTTCAC	3485
Qy	rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr	1037
Db	ATACATGTAGAGAGATMAAAGTCCACAGCGGAATTTATCGGTATGATGGCAATGACAA	3545
Qy	gGluGlyProProLeuCys--ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuPr	1056
Db	ACAAGGACAGATGATCGTGTTCATTGTATAGAACTTTTGAAGAATAATGGAAGATTACC	3605
Qy	oProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTyrPalaProGl	1076
Db	AAGACCAGATGGATGCCCGCAGATGAGATCTATATGATCATGACAGATGCTGGAACAATAA	3665
Qy	uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu	1093
Db	TGTAAATCAACGCCCTCCCTTGGGATCTAGCTCTTCGATGGATCAAAATA	3717

assessing therapy, comprises measuring levels of Stat or their inhibitors -

Example; Fig 8; 218pp; English.

The present invention relates to a method for monitoring acceptance of a transplant or an autoimmune disease in a mammal. The method comprising determining the amount of at least one of Stat4 (signal transducer and activator of transcription), Stat6, SOCS1 (suppressor of cytokine signalling) or SOCS3 mRNA or protein in a sample of the transplant, taken from the host or an affected tissue sample. Stats are activated by receptor-associated Janus kinases (Jaks) which include Jak1, Jak2, Tyk2, Jak3. The method is used to determine whether acceptance of a transplant has been induced or to determine if autoimmune disorders (systemic lupus erythematosus, glomerulonephritis, rheumatoid arthritis, Wegener's granulomatosis, chronic active hepatitis, atopic dermatitis, multiple sclerosis, myasthenia gravis, haemolytic anaemia, nephrotic syndrome, thyroiditis, diabetes mellitus, Crohn's disease, inflammatory bowel disease and vasculitis) are being treated successfully and may be used to adjust treatment regimes. The present sequence is human Jak2 DNA.

Sequence 5117 BP; 1623 A; 927 C; 1103 G; 1464 T; 0 other;

Alignment Scores:	2,15e-190	Length:	5117
Pred. No.:	2621.00	Matches:	533
Score:	66.12%	Conservative:	193
Percent Similarity:	48.54%	Mismatches:	341
Best Local Similarity:	44.73%	Indels:	31
Query Match:	24	Gaps:	14
DB:			
US-09-397-967A-16 (1-1099) x AAD24311 (1-5117)			
Qy	20	SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl	39
Db	585	TCAGTCTTCAGGT-----GTATCTTTACCATTCCTCTGGGAATCTCGAGCAGA	635
Qy	39	nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy	59
Db	636	TTATCTGACCTTCCATCTGGGGAGTATGTTGCAGAGAAATCTGTATTCTGCTCTTAA	695
Qy	59	sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe	79
Db	696	AGCTTGTGGTATCACACCTGTGTATCATATAATGTTTGTAAATGAGTCAACAGAAAG	755
Qy	79	rCysTrpPheProProSerHisIlePheCysIleGluAspValAspThrClnValLeuVa	99
Db	756	GATCTGTATTCACCAACCATCTTCCATATAGATGAGTCAACGAGCATTAATGTACT	815
Qy	99	lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr	117
Db	816	CTACAGAAATGAATTTTACTTTCTCTGTTGATATTCAGTGGCAGCAACAGACCTTATCG	875
Qy	117	gPheGlyLeuArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe	137
Db	876	GCATGGAATATCTCGAGGTCTCGAGCTCCTCTTCTTGATGACTTTTGTCTATCTTACCT	935
Qy	137	uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy	157
Db	936	CTTTGTCTCAGTGGCGCATGATTTGTGTCAGCATGATGATAAAGTACCTGTACTCATGA	995
Qy	157	sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl	177
Db	996	AACACAGGAAGATGTTCTGGGATGGCAGTGTAGATATGATGAGAAATACCCAAAGNAA	1055
Qy	177	nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrIleAlaCysLeuProProSe	197
Db	1056	CGATCAAAACCCACTGGCCATCTATAACTCTATCAGCTACAAGACATTTCTACCAAAATG	1115
Qy	197	rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgGluIleArgThrVa	217
Db	1116	TATTTCGAGCAAGATCCAGACACTATCATATTTTTCACAAAGGACGATATAGGTACAGATT	1175



Db 3324 TCACATAAAACCTTCTCAGTACACATCTCAGATATGCAAGGGTATGAGTATCTTGGTAC 3383  
 Qy 937 aargargCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaH 957  
 Db 3384 AAAAAGGTATATCCACAGGATCTGGCAGCAGAAATATATGTTGGTGAACAGAACAG 3443  
 Qy 957 sValIysIleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLysAspTyrTyrVa 977  
 Db 3444 AGTTAAATTTGGAGATTTGGGTAAACCAAGTCTTCCACAGACAAAGATACATATAA 3503  
 Qy 977 lValArgGluProGlyGluSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs 997  
 Db 3504 AGTAAAGAACCTGGTGAAGTCCCATATTTCTGGTATGCTCCAGAAATCACTGACAGAG 3563  
 Qy 997 nIlePheSerArgGluSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh 1017  
 Db 3564 CAAAGTTTCTGGCTCAGATGTTTGGAGCTTTGGAGTGTCTGTATGAATCTTTCAC 3623  
 Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr 1037  
 Db 3624 ATACATTGAGAAGATAAAGTCCACAGCGGAATTTATGCGTATGATTGGCAATGACAA 3683  
 Qy 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr 1056  
 Db 3684 ACAAGCAGACATGATCGTGTCTCCATTTGTAGAACTTTTCAGAGATAATGGAAGATTACC 3743  
 Qy 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076  
 Db 3744 AAGACGATGATGATGCCACGATGATCTATATGATCATGACAGATGCTGGGAACATAA 3803  
 Qy 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093  
 Db 3804 TGTAATACAGGCCCTCTTTAGGATCTAGCTCTCGAGTGGTCAATAA 3855  
 RESULT 9  
 ID AAQ85412 standard; cDNA; 3629 BP.  
 AC AAQ85412;  
 XX  
 XX 25-MAR-2003 (updated)  
 DT 05-OCT-1995 (first entry)  
 XX  
 XX Murine JAK2 kinase coding sequence and flanking regions.  
 XX  
 XX JAK family; protein tyrosine kinase; cytokine receptor; mouse;  
 KW phosphorylation; signal transduction; activation; ss.  
 XX  
 XX Mus musculus.  
 PH  
 FT CDS Location/Qualifiers  
 FT 94..3483  
 FT /\*tag= a  
 FT /product= Jak2\_kinase  
 FT 554..556  
 FT /\*tag= b  
 FT /note= "Published partial sequence of Jak2 cDNA  
 FT (Harpur et al., Oncogene 7:1347-1353(1992))  
 FT has CCC at this position"  
 FT 1089..3579  
 FT /\*tag= c  
 FT /note= "Published partial sequence of Jak2 cDNA  
 FT (Harpur et al., Oncogene 7:1347-1353(1992))  
 FT differs from AAQ85412 in having the  
 FT nucleotides shown in brackets at the following  
 FT positions: 1089(T), 1103(C), 1114(G), 1119(G),  
 FT 1122(C), 1128(C), 1131(G), 1134(G), 1137(C),  
 FT 1140(G), 1143(G), 1146(C), 1188(T), 1194(G),  
 FT 1230(A), 1245(T), 1260(T), 1266(C), 1272(T),  
 FT 1275(A), 1293(T), 1305(T), 1323(C), 1341(A),  
 FT 1344(A), 1359(G), 1365(A), 1368(T), 1374(T),  
 FT 1401(C), 1413(C), 1431(T), 1458(A), 1476(G),  
 FT 1488(T), 1511..1512(GT), 1578(C), 1590(T),  
 FT

FT 1593(T), 1602(G), 1623(T), 1642(G), 1657(C),  
 FT 1728(G), 1743(C), 1755(C), 1770(A), 1809(G),  
 FT 1816(G), 1821(C), 1857(A), 1878(T), 1935(A),  
 FT 1938(A), 1963(T), 1974(G), 2025(T), 2055(G),  
 FT 2079(C), 2082(C), 2085(C), 2253(A), 2259(G),  
 FT 2283(A), 2285(C), 2433(G), 3453(C), 3579(C)"  
 conflict 2226  
 FT /\*tag= d  
 FT /note= "Location of a 7 amino acid insert detected  
 FT in previous studies, but not in the present  
 FT study"  
 FT 3595..3619  
 FT /\*tag= e  
 FT /note= "Three extra nucleotides (all A's) were  
 FT present in previous studies at positions  
 FT 3595, 3598 and 3619 in the 3'-UTR"  
 XX  
 EN WO9503701-AL.  
 XX  
 XX 09-FEB-1995.  
 PD  
 XX  
 XX 29-JUL-1994; 94WO-US08676.  
 PF  
 XX 29-JUL-1993; 93US-0097997.  
 XX  
 XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.  
 PA  
 XX Ihle JN, Queller FW, Silvennoinen O, Witthuhn BA;  
 PI WPI; 1995-081950/11.  
 XX P-PSDB; AAR70830.  
 DR  
 XX  
 XX Inhibiting a cellular response to a cytokine by inhibiting Jak  
 PT kinase - to treat diseases caused by excessive response to  
 PT cytokine, e.g. erythrocytosis and other cellular proliferative  
 PT diseases  
 XX  
 XX Claim 29; Fig 1; 167pp; English.  
 PS  
 XX Inhibiting the activity of a jak kinase (pref. Jak1, Jak2, Jak3 or  
 CC Tyk2) in a eukaryotic cell is claimed as a method of inhibiting the  
 CC biological response of that cell to a cytokine (not IL-3 or  
 CC erythropoietin). The present sequence (murine JAK2 kinase) encodes a  
 CC 1129 amino acid protein which includes an epitopic sequence at  
 CC amino acid positions 758-776. Antibodies which selectively bind the  
 CC epitope are able to bind Jak2 without interfering with the activity  
 CC of the kinase. Such antibodies are claimed and are useful for  
 CC detecting and extracting Jak2. The 71 nucleotide differences noted  
 CC between the present sequence and the published partial sequence  
 CC result in 9 amino acid changes.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 3629 BP; 1142 A; 733 C; 842 G; 912 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 1,76e-190 Length: 3629  
 Score: 2619.50 Matches: 537  
 Percent Similarity: 65.50% Conservative: 190  
 Best Local Similarity: 48.38% Mismatches: 350  
 Query Match: 44.70% Indels: 33  
 DB: 16 Gaps: 15  
 US-09-397-967A-16 (1-1099) x AAQ85412 (1-3629)  
 Qy 4 ProSerGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla 23  
 Db 145 CCTGTACATCAGAAATGGTGTATTCCTGGAGAGTCTTAATCTGTGACGACATAGACCA 204  
 Qy 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43  
 Db 205 GTCCTTCAAGTGTATCTGTACCATCTCTTGGCAAGCTGAAGGAGAGTATCTGAAGTTT 264  
 Qy 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyTle 63



Db	265	CCAAGTGGAGAGTAGTATGTTGCAGAAAGAAATTTGTGTGGCTGCTTTAAAGACCTGTGGTAT	324
Qy	64	LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTyrPhePro	83
Db	325	ACGCCTGTGTATCATATATGTTTGGCGTATATGAGTGAACCGAAAGAGATCTGTACCCA	384
Qy	84	ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg	103
Db	385	CCCAATCATGTCTTCCCATAGACAGCAACCGCATGACATCTCTACAGGATAGG	444
Qy	104	PheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg	121
Db	445	TTCTACTTCCCTCATATGTAAGTGTAGTGGCAGCAGCAACCTACAGATACGGAGTGTCC	504
Qy	122	LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis	141
Db	505	CGTGGGCTGAAGCTCTCTGCTTGATCACTTTGTCACTGCTTACCTTTTGTCTCAGTGG	564
Qy	142	ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu	161
Db	565	CGGCATGATTTTGTTCACGGATGATAAAAGTACTCTGTGACTCATGAACTCAGGAAGAG	624
Qy	162	PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro	181
Db	625	TGTCCTTGGGATGGCGGTGTAGACATGATGAGAATAGCTAAGGAGAAGACACAGCTCCA	684
Qy	182	GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal	201
Db	685	CTGGCTGCTATAACTCTGTGTCAGTCAAGACATCTTTACCAAGTCGGTTCGAGCGAAG	744
Qy	202	IleGlnGlyGlnAsnPheValThrArgArgArgIleArgThrValValLeuAlaLeu	221
Db	745	ATCCAGACATATCACATTTTAAACCGGAAGCAATCAGGTACAGATTTCGCAGATTCATT	804
Qy	222	LeuProCysGlyArgLeuProGlyArgProTyrLalLeuMetAlaLysTyrIleLeuAsp	241
Db	805	CAGCAATTCAGTCAATGTAAAGCCACTGCCAGGAACCTAAACCTAAGTATCTTTATAAAC	864
Qy	242	LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal-----	256
Db	865	CTGGAAACCTCGAGTCTGGCTTCTACACAGAACAGTTTGAAGTAAAGAAATCTGCAAGA	924
Qy	257	GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle	276
Db	925	GGT--CCITTCAGGTGAGGAGATTTTTCACACCATTTATAAATCTGGAACGGTGAATT	981
Qy	277	ProTyrSerSer-----AsnAspGluLeuPheGlnThr	287
Db	982	CAGTGGTTCAGAGGGAAACATAAGGAAGTGAGACACTGCAGAACAGGACGTACAGTTA	1041
Qy	288	PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro	307
Db	1042	TATTTGTGATTCCTTGATATTTATGATGTGAGTATTAAAGCAGCAACACCG--GAATGC	1098
Qy	308	AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu	327
Db	1099	TCAAAATGAAAGTAGAATGTGTAAGTCTTCCATAAACACAGATGTAAGTTTGGAGATAGAA	1158
Qy	328	PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu	347
Db	1159	CTTAGCTCATATAAAGAACGCTTGTCACTTCGTGTCATTATTCACGGGTATTACAGACTA	1218
Qy	348	IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu	367
Db	1219	ACTGCGGATCGGCACCATTAACCTCTGCAAGAGAGTGGCTCCCCAGCTGTCGAGAAC	1278
Qy	368	GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla	387
Db	1279	ATACACAGCAACTGCCACGCCCAATATCAATGGAATTTTCCATTAGCAACCTAAGAAG	1338
Qy	388	AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe	407

```

QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
Db 2473 TTCAGCCCTGCTTTCAGAGCTGTCATCCGATGCTTTAAACAGCCTGTTTACTCCAGATTAT 2532

QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal 799
Db 2533 GAACACTACTACAGAA---AATGACATGCTACCAACATGAGATAGTAGTCCCTAGGGTTT 2589

QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys 819
Db 2590 TCTGTGTCT-----TTTGAAGACAGGAGCCCTACACAGTTTGAAGAGACACACTTGAAG 2643

QY 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro 839
Db 2644 TTTCTACAGAGCTTTGGCAAGGTAACTTCGGAGTGTGGAGATGCGCCCTATGACCCG 2703

QY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
Db 2704 CTGACAGACAACACTGGCGAGGTGGTGGCTGTGAAGAACTCCAGACACAGCACTGAAGAG 2763

QY 860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879
Db 2764 CACCTCGAGACTTTCAGAGGGAGATCGAGATCTGAAATCTCTGCAGCATGACAACATC 2823

QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
Db 2824 GTCAAGTACAAAGGAGGTGCTCAGTGGCGGTGCGGCAACCTTAAGATTAAATTATGAA 2883

QY 900 TyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg---GlyLeuHisThr 918
Db 2884 TATTACCATATGGAAGTTTACGAGACTATCTCCAAAACATAAAGAACGGATAGATCAC 2943

QY 919 AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 938
Db 2944 AAAAAAATCTTCAATACACATCTCAGATATGCAAGGGCATGGAATATCTTGGTACAAA 3003

QY 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal 958
Db 3004 AGGTATATCACAGGACCTGGCAACAGGAACATATGTGGTGAATGAGAACAGAGGTT 3063

QY 959 LysIleAlaAspPheGlyLeuAlaLysLeuProLeuGlyLysAspTyrTyrValVal 978
Db 3064 AAATAGGAGACTTCGGATTAAACCAAGTCTTGGCGGAGCAAGAAATACTACAAAGTA 3123

QY 979 ArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIle 998
Db 3124 AAGGAGCCAGGGGAAAGCCCATATCTGGTACGACCTGAATCTTTGACGGAGAGCAAG 3183

QY 999 PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3184 TTTTCTGTGGCTCAGATGTGTGGAGCTTTGGAGTGTCTTATACGAACTTTTCACATAC 3243

QY 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038
Db 3244 ATCGAGAGAGTAAAGTCCACCGTGGAATTTATGCAATGATGGAATGATAAACA 3303

QY 1039 GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1057
Db 3304 GGGCAATGATGTGTCTTCCATTTGATAGAGTACTGAAGACCAACGGAAGATTGCCAAG 3363

QY 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluPro 1077
Db 3364 CCAGAAGATGCCAGATGAGATTATGTGATCATGACAGAGTCTGGAACAACAATGTG 3423

QY 1078 HisAspArgProAlaPheAlaThrIleuSer 1087
Db 3424 AGCCAGCGTCCCTCTCTTCAGGGACCTTTCG 3453

RESULT 10
AAC66244
ID AAC66244 standard; cDNA; 3629 BP.
XX

```

```

AC AAC66244;
XX 19-FEB-2001 (first entry)
XX
XX Jak2 polynucleotide sequence.
DE
XX Jak3; kinase; cytokine; cellular response; inhibition; jak2; ss;
KW cell proliferation; erythrocytosis.
XX
XX Unidentified.
OS
XX US6136595-A.
PN
XX 24-OCT-2000.
PD
XX 18-JUN-1996; 96US-0665574.
PF
XX 29-JUL-1994; 94US-0282012.
PR 29-JUL-1993; 93US-0097997.
PR 09-SEP-1993; 93US-0118968.
XX
XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
PA
XX Silvenmoinen O, Witthuhn BA, Ihle J;
PI
XX WPI; 2000-686080/67.
DR
XX P-PSDB; AAB35719.
XX
XX New DNA encoding Jak3 kinase is useful as cytokine regulator for
PT treating cell proliferation -
XX
XX Example 1; Fig 1; 100pp; English.
XX
XX This invention relates to DNA encoding a murine Jak3 protein. The amino
CC acid sequence of the Jak3 protein is given in AAB35715. The Jak family
CC of kinases are involved in the cellular response to the binding of
CC cytokines to their respective receptors. Jak3 kinase mediated activation
CC of some cytokines through their phosphorylation in response to
CC cytokine-receptor binding. Inhibiting the activity of Jak3 kinase (at the
CC nucleic acid level with antisense sequences or ribozymes, or at the
CC protein level with antibodies, kinase inhibitors etc.) is used to treat,
CC or diagnose, diseases caused by excessive secretion of certain cytokines,
CC e.g. excessive cell proliferation such as erythrocytosis. Alternatively,
CC the Jak3 protein and polynucleotide can be used to treat conditions
CC associated with defective Jak3 activity. The DNA sequence can be used to
CC produce recombinant Jak3 and this used to raise antibodies useful as
CC specific inhibitors or to detect or isolate Jak3 without interfering
CC with its enzymatic activity. The present sequence represents cDNA
CC encoding a Jak2 protein used in the isolation and characterisation of the
CC Jak3 protein of the invention.
XX
XX Sequence 3629 BP; 1142 A; 732 C; 843 G; 912 T; 0 other;
SQ

```

```

Alignment Scores:
Pred. No.: 1,76e-190 Length: 3629
Score: 2619.50 Matches: 537
Percent Similarity: 65.50% Conservative: 190
Best Local Similarity: 48.38% Mismatches: 350
Query Match: 44.70% Indels: 33
DB: 21 Gaps: 15

```

```

US-09-397-967A-16 (1-1099) x AAC66244 (1-3629)
QY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerGluAla 23
Db 145 CCTGTACATCAGAAATGGTGTATTTCTTGGAGTCTAATCTGTGACGACATAGAGCCA 204
QY 24 GlyAlaLeuHisValLeuLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43
Db 205 GTCCTTCAAGTGTATCTGTACCATCTCTTGGCAAGCTGAAGGAGATATCTGAAGTTT 264
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValargAlaAlaLysAlaCysGlyTle 63

```



265 CCAAGTGGAGATGTTGTCAGAGAAATTTGTGTGCTGCTCTTAAAGCTTGTGCTATT 324  
64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83  
325 ACGCTGTGTATCAATAATGTTTGGTAAATAGTAGTGAACCCGAAAGGATCTGGTACCCA 384  
84 ProSerHisIlePheCysIleGluLeuValAspThrGlnValLeuValTyrArgLeuArg 103  
395 CCCAATCATGCTCTCCACATAGACGAGTCAACAGGATGACATCTCTACAGGATAGG 444  
104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121  
445 TTCTACTTTCCCTCAITGTGTGTAGTGGCAGCAGAGAACCTACAGATACGAGGTGCC 504  
122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141  
505 COTGGGGTGAAAGCTCTCTGCTTATGACTTTGTATGCTTCTTACCTTTTGTCTCAGTGG 564  
142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161  
565 CGGCATGATTTTGTTCACGGATGGATAAAAGTACCTGTGACTCATGAACCTCAGGAAG 624  
162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181  
625 TGTCTTGGGATGGCGGTGTAGACATGATGAGATAGCTAAGAGAGAACACGAGACTCCA 684  
182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201  
685 CTGGCTGTCTATACTGTCTAGCTTACAGACATCTTACCAAGTGGCTTCGACGGAG 744  
202 IleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAlaLeu 221  
745 ATCCAAGACTATCACATTTTAAACCGGAAGCAATCAGTACAGATTTTCGCAGATTCATT 804  
222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241  
805 CAGCAATTCAGTCAATGTAAAGCCACTGCCAGGAACCTTAAAGTATCTTATAAAC 864  
242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal----- 256  
865 CTGMAACCTTGACTCTGCTCTTACAGAACACAGTTTGAAGTAAAGAAATCTCCAGA 924  
257 GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276  
925 GGT---CCTCAGGTGAGGAGATTTTGCACACATTAATACTGTAACCGTGAAT 981  
277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287  
982 CAGTGGTCAAGAGGGAAACATAAGGAAAGTGAACACTGACAGACACAGGACGTACAGTTA 1041  
288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307  
1042 TATGTGTATTTCCCTGTATTAATGATGTCAGTATTAAGCAAGCAACACG---GAATGC 1098  
308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327  
1099 TCAATGAAGTAGAATGTAACTCTCCATAAACCAAGATGGTAAAGTTTGGAGATAGNA 1158  
328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347  
1159 CTTAGCTCATTAAGAGAGCCCTTGTCAITTCGTGTCTTAATTAATGACGGGTATTACAGACTA 1218  
348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGlu 367  
1219 ACTGGGATGGCCACATTAACCTCTGCAAGAGGTGGCTCCCGACGCTGTCTCAGAAC 1278  
368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIleLysAla 387  
1279 ATACACAGCACTGCCCGCCCAATATCAATGGATTTTGGCATTTAGCAAACTAAGAG 1338  
388 AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspPhe 407  
1339 GCGGGTAAACAGACTGAGCTATATGTCTACGATGCAGCCCTTAAGCACTTCAACAAATAC 1398

408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427  
1399 TTTCTGACCTTTGCTGTGTGAGCGAGAAAATGTCATTGAATATAAACACATGTTTGTATACG 1458  
428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447  
1459 AAGATGAGATGGAGAAATACAACCTCAGCGGGAATAAGAGAACTTCAGTAACCTTAAG 1518  
448 GluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467  
1519 GACCTTTTGAATTTGTACAGATGAAAATGTGCGCTCAGACAGTATCATCTTCAGTTT 1578  
468 ThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgGly 487  
1579 ACCAAATGCTGCCCCCAAGCCAAAGATAAATCAACCTTCTCGCTTTCAGA----- 1632  
488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504  
1633 ACAATGGTATTTTCTGATGTTCTCAGATCTCACCAACATACAGAGGCATAAATATGTGAAT 1692  
505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHis 524  
1693 CAAATGGTGTTCACAAAATCAGGAATGAAGATTTAATATTTAATCAAAAGTCTTGCCCAA 1752  
525 GlySerPheThrLysIlePheArgGlySerArgArgGluValValValAsp---GlyGluThr 543  
1753 GGTACTTTTACAAAATTTTAAAGGTGAAGAGAGATTTGGAGATTTATGGTCAACTG 1812  
544 HisAspSerGluValLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563  
1813 CACAAAACGGAAATCTTTTGAAGTCTTAGATAAGCACATAGGAACATATTCAGAGTCT 1872  
564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583  
1873 TTTCTCGAAGCAGCAAGCATGATGATCAGCTTCTCACAAGCATTTGGTTTGAATAT 1932  
584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602  
1933 GGTGTCTGTCTGTGGAGAGGAGAACATTTCTGTTCAAGAAATTTGTAATAATTTGGATCA 1992  
603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnVal 622  
1993 CTGGATACATACCTCAAGAGAAACAAAATTTCCATAATATATATGGAACCTTGAGTG 2052  
623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsn 642  
2053 GCTAAGCAGTTGGCATGGCCATGCTTTCTAGAAAGAAAATCCCTTATTTCATGGAAAT 2112  
643 ValSerAlaArgLysValLeuAlaArgGluGlyGly-----AspGlyAsnProPro 660  
2113 GTGTGTCTAAATAATCTCTCTATCAGAGAGAAAGACAGAGAACCGGGGAACCCCT 2172  
661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680  
2173 TTCATCAACTTAGTATCTTGGCATTAGCATTACAGTTCTACCGAAGACATCTTCTAG 2232  
681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700  
2233 CAGAGAATACCATGGTACCTCTCTGAATGCATGAGAATCTCAAAAATCTCAATCTGCA 2292  
701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720  
2293 ACAGCAAGTGGAGCTTGGGAGCCACTCTGTGGAGATCTCGAGTGGAGAGATAAGCCC 2352  
721 IleThrSerLeuGluProAlaLysLysLysPheTyrGluAspGlnGlyGlnLeuPro 740  
2353 CTGAGTGTCTGGATTTCTCAAGAGAAAGCTGCGATTTCTATGAAGATAGCATCAGCTTCT 2412  
741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
2413 GCACCAAGTGGACAGAGTTAGCAAACTTATAAATAATTCATGAGTGCATGAGCCAGAT 2472

QY	761	ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr	780	XX	19-JUN-2001	(first entry)	XX
Db	2473	TTTCCAGCCCTTTCAGACTGTCATCCGTGATCTTAAAGCTGTTTACTCCAGATTAT	2532	XX	Murine (Janus kinase 2) Jak2 cDNA.		XX
QY	781	GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal	799	XX	Mouse; Janus kinase 2; Jak2; therapy; cytokine; tyrosine kinase;		XX
Db	2533	GAATCTACTAACAGAA---AATGACATGCTACCAACATGATAGATAGTGCCTAGGCTT	2589	KW	antiproliferative; cyostatic; cell proliferative disorder;		KW
QY	800	AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys	819	XX	cellular response; erythrocytosis; ss.		XX
Db	2590	TCTGTGTCT-----TTTGAAGACAGGGACCTTACACAGTTTGAAGAGACACTTGAAG	2643	OS	Mus sp.		OS
QY	820	TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuLeuCysArgTyrAspPro	839	FH	Key	Location/Qualifiers	FH
Db	2644	TTTCTACAGCAGCTTGGCAAGGTAATCTTCGGAGGTGTGGAGATGCGCTATGACCCG	2703	FT	5'UTR	1..93	FT
QY	840	LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp	859	FT	FT	/*tag= a	FT
Db	2704	CTGCAGGACAACTGGCGAGGTGGTGGCTGTGAAGAACTCCAGCACACGACTGAAGAG	2763	FT	CDS	94..3483	FT
QY	860	GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle	879	FT	FT	/*tag= b	FT
Db	2764	CACCTCCGAGACTTTCAGAGGGAGATCGAGATCTGAAATCCTTCGAGCATGACACATC	2823	FT	misc_feature	523	FT
QY	880	ValLysTyrArgGlyValSerTyrGlyProGlyValGlnSerLeuArgLeuValMetGlu	899	FT	FT	/*tag= c	FT
Db	2824	GTCAGTACAGGAGGTGCTACAGTCGGGTGCGGCAACCTTAGATTAAATTATGGAA	2883	FT	FT	/note= "5, end of published partial Jak2 cDNA	FT
QY	900	TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr	918	FT	FT	replace (551..553, CCC)	FT
Db	2884	TATTTACCATATGGAAGTTTACGAGACTATCTCCAAAACATAAAGAACGATAGATCAC	2943	FT	FT	/*tag= d	FT
QY	919	AspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg	938	FT	FT	replace (1089, T)	FT
Db	2944	AAAAAACTCTTCAATACACATCTTCAGATATGCAAGGGCATGGAATATCTTGTACAAA	3003	FT	FT	/*tag= e	FT
QY	939	ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal	958	FT	FT	replace (1103, C)	FT
Db	3004	AGTATATCCACAGGACCTGGCAACAGGAACATATTGTGGAAATGAGACAGGGTT	3063	FT	FT	/*tag= f	FT
QY	959	LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal	978	FT	FT	replace (1111, G)	FT
Db	3064	AAAAATAGGAGACTTCGGATTAAACCAAAGTCTTCCGCGAGGACAAAGAAATACTACAAAGTA	3123	FT	FT	/*tag= g	FT
QY	979	ArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAsnIle	998	FT	FT	replace (1119, G)	FT
Db	3124	AAGGAGCCAGGGGAAACCCCATATCTGTGACGACCTGAAATCCTTTCGCGAGGACGAA	3183	FT	FT	/*tag= h	FT
QY	999	PheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheThrTyr	1018	FT	FT	replace (1122, C)	FT
Db	3184	TTTTCTGTGGCTCAGATGTGTGGAGCTTTGGAGTGTCTTATACGAACCTTTTCACATAC	3243	FT	FT	/*tag= i	FT
QY	1019	CysAspLysSerCysProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu	1038	FT	FT	replace (1128, C)	FT
Db	3244	ATCGAAGAAGTAAAGTCCACCCGTGGAATTTATGCGAATGATTGCGAATGATAAACA	3303	FT	FT	/*tag= j	FT
QY	1039	GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro	1057	FT	FT	replace (1131, G)	FT
Db	3304	GGGCAATGATGTGTTCATTTGATAGACTCTGAAGGCAACGGAAGATTGCCAAGG	3363	FT	FT	/*tag= k	FT
QY	1058	ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTyrAlaProGluPro	1077	FT	FT	replace (1134, G)	FT
Db	3364	CCAGAAGATGCCAGATGAGATTTATGTGATCATGACAGAGTCTGGAACAACAATGTG	3423	FT	FT	/*tag= l	FT
QY	1078	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1137, C)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= m	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1140, G)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= n	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1143, G)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= o	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1146, C)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= p	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1188, T)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= q	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1194, G)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= r	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1230, G)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= s	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1245, T)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= t	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1260, T)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= u	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1266, C)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= v	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1272, T)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= w	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1275, Q)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= x	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1293, T)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= y	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1305, T)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= z	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1323, C)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= aa	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1341, A)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= ab	FT
QY	1087	HisAspArgProAlaPheAlaThrIleuSer	1087	FT	FT	replace (1344, A)	FT
Db	3424	AGCCAGCGTCCCTCTTCAGGGACCTTTCG	3453	FT	FT	/*tag= ac	FT

RESULT 11  
AAD03607  
ID AAD03607 standard; cDNA; 3629 BP.  
XX  
AC AAD03607;

```

FT conflict replace (1359, G)
FT /tag= ad
FT conflict replace (1365, Q)
FT /tag= ae
FT conflict replace (1368, T)
FT /tag= af
FT conflict replace (1374, T)
FT /tag= ag
FT conflict replace (1401, C)
FT /tag= ah
FT conflict replace (1413, C)
FT /tag= ai
FT conflict replace (1431, T)
FT /tag= aj
FT conflict replace (1453, Q)
FT /tag= ak
FT conflict replace (1476, G)
FT /tag= al
FT conflict replace (1488, T)
FT /tag= am
FT conflict replace (1511..1512, GT)
FT /tag= an
FT conflict replace (1578, C)
FT /tag= ao
FT conflict replace (1590, T)
FT /tag= ap
FT conflict replace (1593, T)
FT /tag= aq
FT conflict replace (1602, G)
FT /tag= ar
FT conflict replace (1623, T)
FT /tag= as
FT conflict replace (1642, G)
FT /tag= at
FT conflict replace (1657, C)
FT /tag= au
FT conflict replace (1728, G)
FT /tag= av
FT conflict replace (1743, C)
FT /tag= aw
FT conflict replace (1755, C)
FT /tag= ax
FT conflict replace (1770, A)
FT /tag= ay
FT conflict replace (1809, G)
FT /tag= az
FT conflict replace (1816, G)
FT /tag= ba
FT conflict replace (1821, C)
FT /tag= bb
FT conflict replace (1857, A)
FT /tag= bc
FT conflict replace (1878, T)
FT /tag= bd
FT conflict replace (1935, A)
FT /tag= be
FT conflict replace (1938, A)
FT /tag= bf
FT conflict replace (1963, T)
FT /tag= bg
FT conflict replace (1974, G)
FT /tag= bh
FT conflict replace (2025, T)
FT /tag= bi
FT conflict replace (2055, G)
FT /tag= bj
FT conflict replace (2079, C)
FT /tag= bk
FT conflict replace (2082, C)
FT /tag= bl
FT misc_feature
FT /tag= bm
FT /note= "Nucleotides corresponding to 7 amino acid insert"
FT

FT conflict replace (2253, A)
FT /tag= bn
FT conflict replace (2259, G)
FT /tag= bo
FT conflict replace (2283, A)
FT /tag= bp
FT conflict replace (2285, C)
FT /tag= bq
FT conflict replace (2433, G)
FT /tag= br
FT conflict replace (3453, C)
FT /tag= bs
FT 3'UTR 3484..3629
FT /tag= bt
FT conflict replace (3579, C)
FT /tag= bu
FT conflict replace (3595, A)
FT /tag= bv
FT conflict replace (3599, A)
FT /tag= bw
FT conflict replace (3620, A)
FT /tag= bx

XX US6210654-B1.
PN
XX
XX 03-APR-2001.
XX
PF 08-OCT-1997; 97US-0946994.
XX
PR 18-JUN-1996; 96US-0665574.
PR 29-JUL-1993; 93US-0097997.
XX
PA (SJUD-) ST JUDE CHILDREN'S HOSPITAL.
XX
PI Ihle J, Witthuhn BA, Quelle FW, Silvennoinen O;
XX WPI: 2001-265367/27.
DR P-PSDB; AAE00352.
XX
XX Modulating a biological response mediated by Jak kinase 2 activation to
PT a cytokine, useful for treating excessive proliferation of eukaryotic
PT cells, comprises inhibiting or enhancing tyrosine kinase activity of
PT Jak kinase in the cell -
XX
PS Example 1; Fig 1; 100pp; English.
XX
CC The present sequence is a cDNA encoding murine (Janus kinase 2) Jak2
CC tyrosine kinase. Jak2 sequence has a 600 amino acid long N-terminus that
CC lacks obvious SH2 (Src homology 2) and SH3 domains. Following this is a
CC kinase related domain (domain 2) and a carboxyl kinase domain (domain
CC 1). Jak kinases mediate cytokine activity through their tyrosine

Alignment Scores:
Pred. No.: 1,76e-190 Length: 3629
Score: 2619.50 Matches: 537
Percent Similarity: 65.50% Conservative: 190
Best local Similarity: 48.38% Mismatches: 350
Query Match: 44.70% Indels: 33
DB: 22 Gaps: 15

US-09-397-967A-16 (1-1099) x AAD03607 (1-3629)
QY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla 23
Db 145 CCTGTACATCAGAAATGGTGATATTCCTGGAAGTGCTAATTCCTGTGAAGCAGATAGAGCA 204
QY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlyProGlnArgLeuSerPhe 43
Db 205 GTCCTTCAAGTGATGTATCTACCATTCCTCTGGCAAGCTGAAGGAGAGTATCTGAAGTTT 264
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLyAlaCysGlyIle 63
Db 265 CCAAGTGGAGAGTAGTTGTGCAGAGAAATTTGTGGCTGCTCTTAAGAGCTGTGGTATT 324

```

QY 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83  
DB 325 AGCCCTGTGTATCATATAATATGTTGGTAAATGAGTGAAACCCAAAGGATCTGGTACCCA 384  
QY 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103  
DB 385 CCCAATCATGCTTCCACATAGACGAGTCAACAGGCGATGACATCTCTACAGGATAAGG 444  
QY 104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121  
DB 445 TTCTACTTCCCTCATTTGGTACTGTAGTGGCAGCAGCAACCTACAGATACGAGGTGCC 504  
QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141  
DB 505 CGTGGGCTCAAGCTCTCTGCTGTGATGACTTGTGTCATGCTTACCTTTTGTCTCAGTGG 564  
QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161  
DB 565 CGGCATGATTTTGTTCACGGATGGATAAAAGTACCTGTGACTCATGAACTCAGGAAGAG 624  
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181  
DB 625 TGTCTTTGGGATGCGGTGTGTAGACATGATGAGAAATAGCTAAGGAGAAAGACCCAGACTCCA 684  
QY 182 GlyGluLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201  
DB 685 CTGGTGTCTATAACTCTGTCTAGCTACAGACATCTTACCAAGTCCGTTCAGCGAAG 744  
QY 202 IleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAlaLeu 221  
DB 745 ATCCAGACTATCACATTTTAAACCGGAGCGAATCAGGTACAGATTTCCGACAGATTCATT 804  
QY 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241  
DB 805 CAGCAATTCAGTCAATGTAAAGCCACTGCCAGGAACCTAAACCTAAGTATCTTATAAAC 864  
QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal-----256  
DB 865 CTGGAAACCTGAGCTGTCTGCTCTACACAGACAGTTTGAAGTAAAGAAATCTGCAAGA 924  
QY 257 GlyLeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276  
DB 925 GGT--CCTTCAGTGGAGAGATTTTGGCAACCATTAATAACTGGAACCGTGGAAATT 981  
QY 277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287  
DB 982 CAGTGTGTCAAAGGGGAAACATAAGAAAGTGAGACACTGACAGAACAGGACGTACAGTTA 1041  
QY 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307  
DB 1042 TATTGTGATTTCCCTGTATATATTGATGTGATGATTAAGCAAGCAAAACCCAG--GAATGC 1098  
QY 308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327  
DB 1099 TCAAAATGAAAGTAGAATTGTAACTGTCCATAAACAAGATGCTAAAGTTTTGGAGATAGAA 1158  
QY 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347  
DB 1159 CTTAGCTCATTAAGAAAGCCCTGTCTCATTTGTGTCATTAATTGACGGGTATTACAGACTA 1218  
QY 348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuGluGlu 367  
DB 1219 ACTCGGATGGCCACATTACTCTGCAAGAGGTGGTCCCCAGCTGTGCTCGAGAAC 1278  
QY 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla 387  
DB 1279 ATACACAGCAACTGCCAGGCCCAATATCAATGGATTTTGGCATTAGCAAACTAAAGAG 1338  
QY 388 AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe 407  
DB 1339 GCGGGTAACAGACTGGGACTATATGTGCTAGCTGACGCCCTTAAGGACTTCAACAAATAC 1398

QY 408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427  
DB 1399 TTTCTGACCTTGTCTGTGCGAGAAAATGTCAATCAATATAAACAACACTGTTTGAATACG 1458  
QY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447  
DB 1459 AAGAATGAGAATGGAGAAATACAACCTCAGCGGACCTAAGAGGAACCTTCAGTAACCTTAAG 1518  
QY 448 GluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467  
DB 1519 GACCTTTTGAATTTGCTACAGATGGAAAATGTGCGCTCAGACAGTATCATCTTCCAGTTT 1578  
QY 468 ThrSerCysAlaProArgProLysGluLysSerAsnLeuIleValValArgArgGly 487  
DB 1579 ACCAAATGTGCTGCCCCCAAGCCAAAAGATAAATCAAACTTCTCGTCTTCAGA-----1632  
QY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504  
DB 1633 ACAATGGTATTTCTGATGTTTCAGATCTCACCAACATTTACAGAGGCATAATAATGTAAT 1692  
QY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHis 524  
DB 1693 CAAATGGTGTTCACAAAATCAGGAATGAAGATTTAATATTAAATGAAGTCTTGGCCAA 1752  
QY 525 GlySerPheThrLysIlePheArgGlySerArgGluValValAsp-----GlyGluThr 543  
DB 1753 GGTACTTTTACAAAAATTTTAAAGGTGTAAAGAGAGAAAGTTGGAGATTATGGTCAAATG 1812  
QY 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563  
DB 1813 CACAAAACGGAGTCTTTTGAAGTCCCTAGATAAAGCACATAGGAACCTATTTCAGAGTCT 1872  
QY 564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583  
DB 1873 TTCTTCCAAACAGCAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932  
QY 584 GlyValCysMetAlaGly--AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602  
DB 1933 GGTGTCTGTCTGTGGAGAGGAGAACATTTCTGTTCAAGAAATTTGTAATAATTTGGATCA 1992  
QY 603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrPheValLeuVal 622  
DB 1993 CTGGATACATACCTGAGAGAGAAACAAATTTCCATAAATATATTATGGAACCTTGGAGTG 2052  
QY 623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsn 642  
DB 2053 GCTAAGCAGTTGGCATGGCCATGTCATTTCTAGAACAAAATCCCTTATTTCATGGGAAT 2112  
QY 643 ValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnProPro 660  
DB 2113 GTGTGTCTAAAAAATATCTGCTTATCAGAGAGAGAGACAGAGGAGAACCGGAGAACCCACCT 2172  
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680  
DB 2173 TTCATCAACTTAGTGTATCTGCGCATTAGCATTTACAGTTCTTACCGAAGGACATTTCTTCAG 2232  
QY 681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700  
DB 2233 GAGAGAATACATGGGTACCTCTCGAATGATGATGAGAAATCCTAAAAATCTCAATCTGCA 2292  
QY 701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720  
DB 2293 ACAGACAAGTGGAGCTTCGGGACCCACTCTGTGGAGATCTGCAAGGAGAGATAAGCC 2352  
QY 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740  
DB 2353 CTGAGTCTCTGGATTTCTCAAGAAAGCTGCGAGTTCTATGAAGATAAGCATCAGCTTCTCT 2412  
QY 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
DB 2413 GCACCAAGTGCACAGAGTTAGCAAACTTTATAAATATTCATGCTGAGCATGAGCCAGAT 2472  
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780

```
Db 2473 TTTCAGCCCTGCTTTTCAGAGCTGTTCATCCGCTGATCTTAACAGCCTGTTTACTCCAGATTAT 2532
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal 799
Db 2533 GAACACTAACAGAA---AATGACATGCTACCAAAACATGATAATAGGTGCTCCAGGGTTT 2589
QY 800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys 819
Db 2590 TCTGGTGCT-----TTTGAAGACAGGACCCTACAGTTTGAAGAGACACTTGAAG 2643
QY 820 TyrIleSerLeuLeuGlyIysGlyAenPheGlySerValGluLeuCysArgTyrAspPro 839
Db 2644 TTCTACAGCAGCTTGCGAAAGTAACCTCGGAGGTGCGAGATGCGCGCTATGACCCG 2703
QY 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
Db 2704 CTGAGGACAACTCTGGCAGGTGGTGTGTGAAGAACTCCAGACAGCACTGAAGAG 2763
QY 860 GlnGlnArgAspPheGlnArgGluIleGlnLeuLysAlaLeuHisSerAspPheIle 879
Db 2764 CACCTCCGAGACTTTGAGAGGAGATCGAGATCTGAATCTTTCGAGCATGACAACATC 2823
QY 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
Db 2824 GTCAAGTACAGGGAGGTGCTACAGTCGGGTGCGCGCAACCTAAGATTAATATGGAA 2883
QY 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr 918
Db 2884 TATTTACCATATGGAAGTTTACGAGACTATCTCAAAAGATATCTCAAAAGACGATAGATCAC 2943
QY 919 AspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 938
Db 2944 AAAAAACTTCTCAATACATCTCAGATATGCAAGGGATGGAATATCTTGGTACAAA 3003
QY 939 ArgCysValHisArgAspLeuAlaArgAsnIleLeuValGluSerGluAlaHisVal 958
Db 3004 AGGTATATCCAGGGACCTGGCAACAGGAACATATTGGTGAAATGAGAACAGGGTT 3063
QY 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978
Db 3064 AAAATAGGAGACTTCGATATACCAAAAGTCTTGGCGCAGCAAAAGAACTACAAAGTA 3123
QY 979 ArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIle 998
Db 3124 AAGGAGCCAGGGGAAAGCCCATATTCTGTGACGCACTGAATCTTTCGAGAGCAAG 3183
QY 999 PheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
Db 3184 TTTCCTGTGGCCTCAGATGTGTGGAGCTTTGGAGTGGTTCTATACGAACTTTTCATAC 3243
QY 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetClyProGluArgGlu 1038
Db 3244 ATCGAGAGAGTAAGATCCACCGTGGAAATTTATCGGAATGATGGCAATATAAACAA 3303
QY 1039 GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgLeuProPro 1057
Db 3304 GGGCAATATGATTGTCTTCATTGTAGAGCTACTGAAGAGCAACCGAAGATTGCCAAG 3363
QY 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluPro 1077
Db 3364 CCAGAAGGATGCCAGATGAGATTTATGTGATCATGACAGAGTGTCTGGAACAACAATGTG 3423
QY 1078 HisAspArgProAlaPheAlaThrLeuSer 1087
Db 3424 AGCCAGCGTCCCTCCTTCAGGACCTTCG 3453

RESULT 12
ABQ76194
ID ABQ76194 standard; DNA; 3503 BP.
XX
AC ABQ76194;
XX
```

```
DT 18-OCT-2002 (first entry)
DE Human Jak2 cDNA.
XX
XX
KW Jak2 kinase; human; immunosuppressive; cytostatic; chromosome 9p23-24;
KW autoimmune disorder; malignancy; leukaemia; gene therapy; vaccine;
KW Jak2; ss.
XX
OS Homo sapiens.
XX
PN CA2203706-A.
XX
PD 25-OCT-1998.
XX
XX
PF 25-APR-1997; 97CA-2203706.
XX
PR 25-APR-1997; 97CA-2203706.
XX
PA (HSCR-) HSC RES & DEV LP.
XX
PI Roifman CM;
XX
DR WPI; 2000-024038/03.
DR P-PSDB; ABB84401.
XX
XX
PT Isolated nucleic acid encoding a human kinase, useful for treating
XX autoimmune disorders and malignancies -
XX
XX Disclosure; Page 44-46; 57pp; English.
CC This invention describes a novel isolated nucleic acid encoding a human
CC Jak2 kinase which has immunosuppressive and cytostatic activity. The Jak2
CC kinase of the invention is located on chromosome 9p23-24 and is expressed
CC highly in spleen, lymph nodes and peripheral blood lymphocytes.
CC Expression of the Jak2 polypeptide increases dramatically on activation
CC of mature B lymphocytes. Jak2 and its mutants are useful for treating
CC autoimmune disorders and malignancies such as leukaemia and can be used
CC for gene therapy or vaccine production. This sequence encodes the human
CC Jak2 protein described in the method of the invention.
CC Note: This specification contained no claims.
XX
SQ Sequence 3503 BP; 1180 A; 612 C; 741 G; 970 T; 0 other;

Alignment Scores:
Pred. No.: 4,34e-188 Length: 3503
Score: 2588.00 Matches: 526
Percent Similarity: 66.27% Conservative: 191
Best Local Similarity: 48.61% Mismatches: 335
Query Match: 44.16% Indels: 30
DB: 21 Gaps: 15

US-09-397-967A-16 (1-1099) x ABQ76194 (1-3503)
QY 36 GlyProProGlnArg-----LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeu 53
Db 142 GGGAACTTGAGGACAGATTATCTGACCTTTCATCTGGGAGTATGTTGCAGAAAGAAATC 201
QY 54 CysValArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeu 73
Db 202 TGTATTGCTGCTTCTAAAGCTTGTGTATCACACCTGTGTATCAATAATATGTTGCTTTA 261
QY 74 AlaThrGluAspPheSerCysTrpPheProSerHisIlePheCysIleGluAspVal 93
Db 262 ATGAGTGAACAGAAAGGATCTGGTATCCACCCACCATGCTTCCATATATAGATGATCA 321
QY 94 AspThrGlnValLeuValTyrArgLeuArgPheTyrPheProAspTrpPhe-----Gly 111
Db 322 ACCAGGCATATGTTACTCTACAGATAAGATTTTACTTTCCTCGTTGTTGATTTGCAGTGC 381
QY 112 LeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeu 131
Db 382 AGCAACAGAGCCTATCGGCATGGAATATCTCGAGGTGCTCGAGCTCCTCTTCTTGTATGAC 441
```

Qy	132	HisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuPro	151
Db	442	TTTGTGATGCTTACCTCTTTGCTCAGTGGCGCATGATCTTGTGCACGGATGGATAAAA	501
Qy	152	ValGlyLeuSerMetIysGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAla	171
Db	502	GTACCTGTGACTCGTGAACACAGGAAGAATGCTTGGGAGCGCAGGTGTACATATGATG	561
Qy	172	GlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyIys	191
Db	562	AGATAGCCAAAGAAACGATCAACCCCACTGGCCATCTATAACTCTATCAGCTACAG	621
Qy	192	AlaCysLeuProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArg	211
Db	622	ACATCTTACC AAAAATGTATTTCGAGCAAGATCCAGACTATCATATTTTGA CAAGAA	681
Qy	212	ArgIleArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgPro	231
Db	682	CGAATAAGGTACAGATTTCGCAGATTATTTCAGCAATTTCAGCAATTCAGCAATTC	741
Qy	232	TyrAlaLeuMetAlaLysTyIleLeuAspLeuGluArgLeuHisProAlaAlaThrThr	251
Db	742	AGAAACTTGAAACTTAAGTATCTATTAAATCTGAAACTCTGCAGCTGCGCTTCTACA	801
Qy	252	GluThrPheArgValGlyLeuProGlyAla-----GlnGluGluProGlyLeu	267
Db	802	GAGAAATTTGAAGTAAAAAGAACCTTGAAGTGGTCTCTTCAGTGGAGGAGATTTTTC	861
Qy	268	LeuArgValAlaGlyAspAsnGlyIleProThrSerSer-----	280
Db	862	ATTATAATAACTGGAACCGTGGAAATTCAGTGGTCAAGAGGGAACACATNAAGAAAG	921
Qy	281	-----AsnAspGluLeuPheGluThrPheCysAspPheProGluIleValAspValSer	298
Db	922	ACACTGACAGACGAGATTACAGTTATATTCGATTTTCTATAATATATGATGTCAGT	981
Qy	299	IleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArg	318
Db	982	ATTAAGCAAGCAAAACCAAGAGGGT---TCAAAATGAAAGCGGAGTTGTAACATATCC	1038
Qy	319	MetAspGlyHisIleLeuGlu---AlaGluPheProGlyLeuProGluAlaLeuSerPhe	337
Db	1039	CAAGATGTGTAAAAATCTCGAAATATTGAACTTAGCTCATTAAGGGAAGCTTTGTCTTC	1098
Qy	338	ValAlaLeuValAspGlyTyIlePheArgLeuIleCysAspSerArgHisTyIlePheCysLys	357
Db	1099	GTGTCATTAAATTGATGGATATTTATAGATTAACTGCAGATGCACATCATTAACCTCTG	1158
Qy	358	GluValAlaProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThr	377
Db	1159	GAAGTAGCACCTCCAGCCGTGCTTGAAATAATACAAAGCAACTGTCATGGCCCAATTCG	1218
Qy	378	LeuAspPheAlaIleHisIysLeuIysAlaIleGlySerLeuProGlyThrTyIleLeu	397
Db	1219	ATGTTTGTTCATTAAGTAAACTGGAAGAACCGAGTAACTAGACTGACGTGATGTACTA	1278
Qy	398	ArgArgSerProGlnAspTyIleAspSerPheLeuLeuThrAlaCysValGlnThrProLeu	417
Db	1279	CGATCGATCCGAAGAGCTTTAATAATATTTTCTGACCTTTCGTGTCGAGCGAGAAAT	1338
Qy	418	GlyProAspTyIleGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuVal	437
Db	1339	GTCAATTGAATAAACAACCTGTTTGATTACAAAAAATGAGATGAAGAGTACAACCTCAGT	1398
Qy	438	GlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGly	457
Db	1399	GGGACAAAAGAAACCTTCAGCAGCTCTTAAGATCTTTTGAATCTTTACCAGATGAAACT	1458
Qy	458	LeuArgValAspGlyAlaAlaLeuTyIleThrSerCysCysAlaProArgProLysGly	477
Db	1459	GTTCGCTCAGACAAATANAATTTTCCAGTTTACTAAATCTCTCCCCCAAGCCAAAGAT	1518
Qy	478	LysSerAsnLeuIleValValArg---ArgGlyCysAsnProAlaProAlaProGlyCys	496

Db	1519	AAATCAAAACCTTTAGTCTTCAGAACGAATGGTGTCTCTGATGTACCAACCTCACCAACA	1578
Qy	497	SerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeu	516
Db	1579	TTACAGAGGCGCTACTCATATGAACCAAAATGGTGTCTTCACAAAATCAGAAATGAAGATTG	1638
Qy	517	GluTrpHisGluAsnLeuGlyHisGlySerPheThrIlysIlePheArgGlySerArgArg	536
Db	1639	ATATTTAATGAAGCCCTGGCCAGGCACTTTTACAAAGATTTTAAAGCGGTACCGAGA	1698
Qy	537	GluValValAsp--GlyGluThrHisAspSerGluValLeuLeuLeuLysValMetAspSer	555
Db	1699	GAAGTAGGAGACTACGGTCAACTGCATGAACACAGAGTCTCTTTAAAGTTCGGATAAA	1758
Qy	556	ArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSer	575
Db	1759	GCACACAGAACTATTCCAGAGTCTTCTTTGAAGCAGCAAGTAGATGATGACCAAGCTTCT	1818
Qy	576	TyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAsp---SerIleMetVal	594
Db	1819	CACAAGCATTTGGTTTTAAATTTATGGAGTATGTCTGTGGAGACGAGAATATTATGGTT	1878
Qy	595	GlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuVal	614
Db	1879	CAGAGATTGTAAATTTGGATCACTAGATACATATCTGAAAAAGAAATAAAATTTGTATA	1938
Qy	615	SerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGlu	634
Db	1939	AAATATATTGAAACCTTGAAGTGTCTAAACAGTTGGCATGGGCCATGTCATTTCTAGAA	1998
Qy	635	AspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly	654
Db	1999	GAATAACACCTTATTCATGGGAATGTATGTGCCAAAAATATTCAGCTTTATCAGAGAAGAA	2058
Qy	655	Gly-----AspGlyAsnProPheIleLysLeuSerAspProGlyValSerProThr	672
Db	2059	GACAGGAAGACAGGAACCTCTCTTTCATCAAACTTAGTGACCTGGCATTAGTATTACA	2118
Qy	673	ValLeuSerLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGln	692
Db	2119	GTTTGGCAAGGACATCTTCAGGAGAGATACCATGGGTACCACCTGAATGCATTGAA	2178
Qy	693	GluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGlu	712
Db	2179	AATCCTAAAAATTTAAATTTGGCAACACAGACAATGGAGTTTGGTACAACCTCTCTGGGAA	2238
Qy	713	ValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLeuLysPhe	732
Db	2239	ATCTGCAGTGGAGGAGATAAACCTCTTAAGTGCTCTGGATTCTCAAGAAAGCTCAAAATT	2298
Qy	733	TyrGluAspGlnGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThr	752
Db	2299	TATGAAGATAGCATCAGCTCTTCGCCAAAGTGGGCAGAAATTAGCAAACTTATAAAT	2358
Qy	753	GlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeu	772
Db	2359	AAATTGATGGATTATGAACAGATTTACGGCTCTCTTTCAGAGCATCATACAGAGATCTT	2418
Qy	773	AsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSer	792
Db	2419	AACAGCTTTGTTTACTCCAGATTATGAATCTTTAAACAGAA---ATGACATGTTACCAAA	2475
Qy	793	ProArgAsp--GluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAla	811
Db	2476	ATGAGAAATGGCCCTAGGTTTCTGGTCCC-----TTTGAAGACCGGATCCTACA	2522
Qy	812	IlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySer	831
Db	2530	CAGTTTGAAGAGAGACATTTGAAATTTCTACAGCAACCTTGGCAAGGGTAAATTTTGGGAGT	2589
Qy	832	ValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLys	851

Db 2590 GTGGAGTGTGGCGGTATGACCTCTACAGACAACTGGGAGGTGGTGGCTGTAAAA 2649

Qy 852 GlnLeuGlnHisSerValProAspGlnGlnArgCysPheGlnArgGluLeuGlnLeu 871

Db 2650 AGCTTCACATAGTACTGAGAGCACCTTAAGAGACTTTGAAAGGAAATTTGAAATCCCTG 2709

Qy 872 LysAlaLeuHisSerAspPheLeuValLysTyrArgGlyValSerTyrGlyProGlyArg 891

Db 2710 AAATCCCTACAGATGACACACATTTGAAAGTACAGGAGTGTGTACAGTGTGTCTGG 2769

Qy 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911

Db 2770 CCGAATCTTAAGTTAATTTGGAATATTTACCATATGGAAGTTTACGAGATATCTTCAA 2829

Qy 912 ArgHisArg--GlyLeuHisThrAspArgLeuLeuPheAlaTrpGlnLeuCysLys 930

Db 2830 AAACATAAGACGGATAGATCACATAAACTTCTGCAGTACACATCTCAGATATCAAG 2889

Qy 931 GlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnile 950

Db 2890 GGTATGGTGTATCTTGGTACAAAAGGTATATCCACAGGATCTGGCAACGAGAATATA 2949

Qy 951 LeuValGlnSerGluAlaHisValLysLeuAlaAspPheGlyLeuAlaLysLeuPro 970

Db 2950 TTGGTGGAGAACGAGAACAGAGTAAATTTGGGTTAAACCAAGTCTTGCCA 3009

Qy 971 LeuGlyLysAspTyrValValArgGluProGlyGlnSerProIlePheTrpValAla 990

Db 3010 CAAGCAAGAATACATAAAGTAAAGAACCTGGTAAAGTCCCATATCTGGTATGCT 3069

Qy 991 ProGluSerLeuSerAspAsnilePheSerArgGlnSerAspValTrpSerPheGlyVal 1010

Db 3070 CCAGAATCACTCACAGAGAGCAAGTTTCTGTGGCCTCAGATGTTTGGAGCTTTGGAGTG 3129

Qy 1011 ValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeu 1030

Db 3130 GTTCTGTATGAATTTTTCATACATCAATTGAGAAGATAAAAGTCCACACGCGGAATTTATG 3189

Qy 1031 ArgMetMetGlyProGluArgGluGlyProProLeuCys--ArgLeuLeuGluLeuLeu 1049

Db 3190 CGTATGATGGCAATGACAAACAGGACAGATGATGCTGCTTCATTTGATGAACCTTTTG 3249

Qy 1050 AlaGluGlyArgArgLeuProProProThrCysProThrGluValGlnGluLeuMet 1069

Db 3250 AAGAATAATGGAAGATTACCAAGACAGATGGATGCCAGATGAGATCTATATGATCATG 3309

Qy 1070 GlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGln 1089

Db 3310 ACAGATGCTGGACAAATAATGTAATCAACGCGCCCTCTTTAGGGATCTAGTCTCTCGA 3369

Qy 1090 LeuAsp 1091

Db 3370 GTGGAT 3375

RESULT 13

ACA55699

ID ACA55699 standard; cDNA; 3900 BP.

XX AC ACA55699;

AC ACA55699;

XX 06-JUN-2003 (first entry)

DE Human signalling pathway polynucleotide probe SEQ ID NO 297.

XX Human; probe; ss; array element; Parkinson's disease;

KW signalling pathway population; cancer; adenocarcinoma; leukaemia;

XX immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.

OS Homo sapiens.

XX US6500938-B1.

PN 31-DEC-2002.

XX 31-DEC-2002.

PD

XX 30-JAN-1998; 98US-0016434.

XX 30-JAN-1998; 98US-0016434.

XX (INCY-) INCYTE GENOMICS INC.

XX Au-Young J, Seilhamer JJ;

XX WPI; 2003-352189/33.

XX Combination of polynucleotide probes, useful as array elements in a

PT microarray for monitoring the expression of a number of target

PT polynucleotides -

XX Claim 1; SEQ ID NO 297; 65pp; English.

XX The invention relates to a combination which, comprises a number of

CC polynucleotide probes comprising a sequence selected from one of the 1490

CC sequences mentioned in the specification. The combination is useful as an

CC array element in a microarray for monitoring the expression of a number

CC of target polynucleotides. The microarray is particularly useful in the

CC diagnosis and treatment of cancer and immunopathology and neuropathology.

CC The microarray is useful in diagnostics and treatment regimens, drug

CC discovery and development, toxicological and carcinogenicity studies,

CC forensics and pharmacogenomics. The microarray is also useful for

CC monitoring progression of diseases and for developing sophisticated

CC profiles for the effects of currently available therapeutic drugs. The

CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs

CC and genomic fragments and in research and diagnostic applications. The

CC array can detect changes in expression in a large number of genes coding

CC for different signaling pathway populations which can be used to diagnose

CC various diseases including cancer e.g. adenocarcinoma and leukaemia,

CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease

CC and Parkinson's disease. The present sequence represents a polynucleotide

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification but was obtained in electronic format directly from USPTO

CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.

XX

SEQ Sequence 3900 BP; 1267 A; 714 C; 853 G; 1048 T; 18 other;

Alignment Scores:

Pred. No.: 2,968-180 Length: 3900

Score: 2486.50 Matches: 509

Percent Similarity: 64.83% Conservative: 184

Best Local Similarity: 47.61% Mismatches: 329

Query Match: 42.43% Indels: 48

DB: 25 Gaps: 14

US-09-397-967A-16 (1-1099) x ACA55699 (1-3900)

Qy 48 TyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIleLeuProValTyr 67

Db 463 TATCTTACCATTCCCTTTGATTGCTGCTTCAAGCTTGGTATCACACTGTAT 522

Qy 68 HisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPheProProSerHisIle 87

Db 523 CATAATATGTTTGTCTTTAATGAGTGAACAGAAAGGATCTGTATCCACCAACCATGTC 582

Qy 88 PheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArgPheTyrPhePro 107

Db 593 TTCCATATAGTAGTCAACCCAGGCATATATGATCTCTACAGAAATAGATTTTACTTTCT 642

Qy 108 AspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArgLysAspLeuThr 125

Db 643 CG-TGTTATTCAGCTGGGAGCAACAGACCTATCGGCATGGAATATCTCGAGGTCTCAA 701

Qy 126 SerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeu 145

Db 702 GCTCCTCTTGTGATGACTTGTTCATGCTTACCTTTTGTCTCAGTGGCGCATGATTTT 761

Qy 146 ValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeu 165



Db	762	GTGCATGGATGGATAAAGTACCTGCTGACTCATGAACACAGGAAGTATGCTTTGGGATG	821	Db	1839	CCAACTCCACCAACATTACAGAGCCCTACTCATATGAACCAAAATGGTGTTCACAAAATC	1898
Qy	166	AlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeu	195	Qy	512	ProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePhe	531
Db	822	ACAGTGTAGATGATGATAGCAATAGCCAAAGAAAACGATCAAAACCCCGCCATCAT	891	Db	1899	AGAAATGAAGATTGATATTATTAAGAGCCTTGGCCAGGCACCTTTTCAAAAGATTTT	1958
Qy	186	LysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGln	205	Qy	532	ArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeuLeu	550
Db	882	AACTCTATCAGCTACAGACATCTTACCAACATGATTTCGACAAAGATCCAAAGACTAT	941	Db	1959	AAAGGCCTACGAAGAGAGTAGGAGACTACGGTCAACTGCATGCAACAGAAAGTCTTTTA	2018
Qy	206	AsnPheValThrArgArgIleArgThrValValLeuAlaLeuLeuProCysGly	225	Qy	551	LysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeu	570
Db	942	CATATTTCAGAGGAAGCGAATAAGGTACAGATTTCGCAGATTATTCAGCAATTCAGC	1001	Db	2019	AAAGTTCTGGATAAAGCACACAGCAACTATTCAGAGCTTTTCTTGAAGCAGCAAGTATG	2078
Qy	226	ArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeu	245	Qy	571	MetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAsp	590
Db	1002	CAATGCAAAACCCACTGCCAGAACTTGAACCTTAAGTATCTTATAAATCTGGAAACTCTG	1061	Db	2079	ATGAGCAGCTTCTCAACAGCATTGGTTTTAAATATGAGCTATGCTGTGGAGAC	2138
Qy	246	HisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAla	261	Qy	591	---SerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLys	609
Db	1062	CAGTCTGCTTCTACACAGAGAAATTTGAAGTAAAGAACTCGAAGTGTCTTCAGT	1121	Db	2139	GAGAAATATTCTGGTTCAGGAGTTTGTAAAAATTTGGATCCTAGATACATATCTGAAAAG	2198
Qy	262	GlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer---	280	Qy	610	ArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAla	629
Db	1122	GAGGAGATTTTGCACCACTTATAATAACTGGAACCGTGAATTCAGTGTGCAAGAGGG	1181	Db	2199	AATAAAATTTGTATAAATATATTATGAAACTTTGAAGTGTCTTAACAGTTGGCAATGCC	2258
Qy	281	-----AsnAspGluLeuPheGlnThrPheCysAspPhePro	292	Qy	630	LeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeu	649
Db	1182	AAACATAAAGAAAGTGAGACACTGACAGAACAGGATTTACAGTTATATTGCGATTTTCT	1241	Db	2259	ATGCAATTTCTAGNAGAAAACACCCCTATTATCATGGGATGTATGTGCCAAAAATATTCTG	2318
Qy	293	GluIleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArg	312	Qy	650	LeuAlaArgGluGlyGly-----AspGlyAsnProProPheIleLysLeuSerAspPro	667
Db	1242	AAATATTATGATGTCAGTATTTAACCAAGCAACCAAGAGGT---TCAAATGAAGCCGA	1298	Db	2319	CTTATCAGAGAAGAACAGACAGAGAAATCCTCTTTTCATCAAACTTAGTGATCCT	2378
Qy	313	LeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuPro	332	Qy	668	GlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpValAla	687
Db	1299	GTGTAACTATCCATAAGCAAGATGGTAAATAATCTGGAATTTGAATCTAGCTAATAGG	1358	Db	2379	GGCATTAGTATTACAGTTTTTGGCCAAAGGACATTTCCAGGAGAAATACCATGGGTACCA	2438
Qy	333	GluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArg	352	Qy	688	ProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTyrGlyPheGly	707
Db	1359	GAAGCTTTGCTTTTCGTGTCATTAATGTAGTGATATTATAGATTAACTGCGAGATGCAT	1418	Db	2439	CCTGAATGCATTGAAAATCTTAAATTTTAAATTTGGCAACAGACAAATGGAGTTTGGT	2498
Qy	353	HisTyrPheCysLysGluValAlaProProArgLeuGluGluAlaAspValCys	372	Qy	708	AlaThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAla	727
Db	1419	CATTACCTCTGTAAGAAGTAGACCTCCAGCCGTGCTTGAATAATATACAAAGCAACTGT	1478	Db	2499	ACCACTTTGTGGGAAATCTCGAGTGAGAGATAAAGCTCTCTAGTGTCTGGATTCTCAA	2558
Qy	373	HisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuPro	392	Qy	728	LysLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeu	747
Db	1479	CATGCCCAATTTTCGATGATTTTGGCATTAGTAACTGAAGAACAGAGGTAAATCAGACT	1538	Db	2559	AGAAAGCTCAAAATTTATGAAGATAGGATCAGCTTCTCGCCAAAGTGGGCAATTA	2618
Qy	393	GlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCys	412	Qy	748	AlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAla	767
Db	1539	GGACTGTATGATCTCGATCGAGCTCTAAGGACTTTAATAAATATTTTTCAGCTTTGCT	1598	Db	2619	GCAACCTTATAAATAATTGTATGATATGAACACAGATTTTCAGGCCTTCTTTTCAGACC	2678
Qy	413	ValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGly	432	Qy	768	IleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspProThr	787
Db	1599	GTCAGCGCAGAAATGTCATTGAATATATAACACTGTTTGATTAACAAAAATGAGATGAA	1658	Db	2679	ATCATACAGATCTTAAACAGTTTGTACTCTCAGATTATGAACATTATAACAGAA---AAT	2735
Qy	433	AlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAla	452	Qy	788	ProGlyIleProSerProArg---AspGluLeuCysValAlaGlyAlaGlnLeuTyrAla	806
Db	1659	GAGTACAACCTCAGTGGGACAAAGAAAGAACTTCAGCAGCTTTAAGATCTTTTGAATTTGT	1718	Db	2736	GACATGTTACCAATATGAGATAGGTGCTTGGGGTTTCTGGTGCC-----TTTGAA	2789
Qy	453	CysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAla	472	Qy	807	CysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLys	826
Db	1719	TACAGATGGAAACTGTTGCTGCATCAGACCAATATAATTTCCAGTTTACTAATGCTGCTCC	1778	Db	2790	GACCGGATCTTACACAGTTTGAAGAGAGACATTGAAATTTCTACAGCAACTTGGCAAG	2849
Qy	473	ProArgProLysGluLysSerAsnLeuIleValValArg---ArgGlyCysAsnProAla	491	Qy	827	GlyAsnPheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyPro	846
Db	1779	CCAAAGCCAAAAGATAAATCAACCTTCTAGTCTTCAGACGAATGGTGTTCATGATGA	1838	Db	2850	GGTAATTTTGGAGTGTGNAGATGTCGGGTATGACCTCTACAGCAACACATGGGCGAG	2909
Qy	492	ProAlaProGlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIle	511	Qy	847	LeuValAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArg	866
		---		Db	2910	GTGGTCTGTGTAATAAAGCTTTCAGCATAGTACTAGAGAGCACCTTAAGAGACTTTGAAAGG	2969



QY 867 GluLeuGlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSer 886  
 DB 2970 GAATTTGAATCCTGAAATCCCTCAGCATGCAACATTTGTAAGTACAGGGATGTGC 3029  
 QY 887 TyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeu 906  
 DB 3030 TACAGTGTGTCGGCGTAATCTAAATTAATTAATGGAATATTTACCATATGGAAGTTTA 3089  
 QY 907 ArgAspLeuLeuGlnArgHisArg--GlyLeuHisThrAspArgLeuLeuPheAla 925  
 DB 3090 CGAGACTATCTTCAAAAAACATAAAGACGATAGATACATAAAATCTTCTGCAGTACACA 3149  
 QY 926 TrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeu 945  
 DB 3150 TCTCAGATATGCAAGGTATGAGATATCTTGTACAAAAGATATATCCACAGGATCTG 3209  
 QY 946 AlaAlaArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeu 965  
 DB 3210 GCAACGAGAAATATATTGTGGAGAACGAGACAGATTTAAATTTGGRGATTTTGGGTTA 3269  
 QY 966 AlalysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerPro 985  
 DB 3270 ACCAAAGTCTTGCCCAAGACAAAGAATACTATAAAGTAAAGAAACCTGGTGAAAGTCCC 3329  
 QY 986 IlePheTrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspVal 1005  
 DB 3330 ATATTCTGTATGTCCTCCAGATCATCTGACAGAGACCAAGTTTCTGTGGCCTCAGATGT 3389  
 QY 1006 TrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerPro 1025  
 DB 3390 TGGACCTTTGGAGTGGTCTGTGATGAACCTTTAC----- 3422  
 QY 1026 SerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCys---Arg 1044  
 DB 3423 -----AAACAGGACAGATCATGTGTTCCAT 3449  
 QY 1045 LeuLeuGluLeuLeuAlaGluGlyArgArgLeuProProProThrCysProThrGlu 1064  
 DB 3450 TTGATAGAACTTTTGAAGAATAATGAGATATACCAAGACCAGATGGATGCCAGATGAG 3509  
 QY 1065 ValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAla 1084  
 DB 3510 ATCTATATGATCATGACAGATGCTGGACAAATAATGTAAATCAACGCCCTCCTTTAGG 3569  
 QY 1085 ThrLeuSerProGlnLeuAspProLeu 1093  
 DB 3570 GATCTAGCTCTTCGAGTGGATCAATA 3596  
 RESULT 14  
 ID AAQ25307  
 AC AAQ25307;  
 XX AAQ25307;  
 XX 25-MAR-2003 (updated)  
 DT 18-FEB-1999 (first entry)  
 XX JAK2 encoding DNA.  
 DE Phosphorylation; JAK1, JAK2; protein tyrosine kinase; human;  
 KW catalytic domain; SH2 domain; growth factor receptor; PTX; murine; ss.  
 OS Mus musculus.  
 XX Key Location/Qualifiers  
 FT 1..2961  
 CDS /\*tag= a  
 FT /product= JAK2  
 FT  
 XX WO9210519-Al.  
 PN  
 XX 25-JUN-1992.

XX 26-NOV-1991; 91WO-US08889.  
 XX 28-NOV-1990; 90AU-0003594.  
 XX (LUDW-) LUDWIG INST CANCER RES.  
 PA Harpur A, Wilks AF, Ziemiecki A;  
 PI WPI; 1992-234591/28.  
 DR P-PSDB; AAR25141.  
 XX Novel protein tyrosine kinase mol. - comprises multiple catalytic  
 PT domains but no SH2 domain and is for phosphorylation of proteins  
 XX Claim 10; Fig 8; 50pp; English.  
 CC This sequence encodes the murine protein tyrosine kinase JAK2 (from  
 CC Janus Kinase). Northern analysis of JAK2 expression in a mouse  
 CC demonstrated the presence of two mRNA transcripts (4.8 and 4.4 kb).  
 CC The levels of these transcripts alter with respect to one another in  
 CC different tissues. The kidney, spleen and lung appear to express  
 CC predominantly the larger form, whereas ovary, placenta, skeletal muscle  
 CC and all murine cell lines analysed express both forms at equal levels.  
 CC The difference in sizes may be due to differential polyadenylation  
 CC sites. Both JAK2 and JAK1 are examples of a new subfamily or class  
 CC of protein tyrosine kinase. These can be used in the phosphorylation  
 CC of proteins, incorporation of labels and in the design of analogues,  
 CC antagonists and agonists of JAK's.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX Sequence 3473 BP; 1128 A; 677 C; 781 G; 887 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 6,71e-172 Length: 3473  
 Score: 2376.50 Matches: 489  
 Percent Similarity: 66.57% Conservative: 166  
 Best Local Similarity: 49.70% Mismatches: 298  
 Query Match: 40.55% Indels: 31  
 DB: 13 Gaps: 14  
 US-09-397-967A-16 (1-1099) x AAQ25307 (1-3473)  
 QY 128 IleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuValSer 147  
 DB 1 CTGCTTGATGACTTTGTCTCATGTTTACCTTTTGTTCAGTGGCGCATGTTTGTCCAC 60  
 QY 148 GlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaVal 167  
 DB 61 GGATGGATAAAAGTACTCTGTGACTCATGAACACTCAGGAAGAGTGTCTTGGGATGGCGTG 120  
 QY 168 LeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThr 187  
 DB 121 TTAGACATGATGAGAATAGCTAAGGAGAAAGACCACTCCACTGGCTGTCTATAACTCT 180  
 QY 188 ValSerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPhe 207  
 DB 181 GTCAGCTACAAGACATTTCTTACCAAGTGGTTCGAGCGAAGATCCAAAGACTATCACCATT 240  
 QY 208 ValThrArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeu 227  
 DB 241 TTAACCGGAAGCGNATCAGGTACAGATTTTCGAGATTCATTCAGCAATTCAGTCAATGT 300  
 QY 228 ProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisPro 247  
 DB 301 AAAGCCACTGCCAGGAACCTTAAACTTAAAGTATCTTATAAACCTGGAACCCCTGCAGTCT 360  
 QY 248 AlaAlaThrThrGluThrPheArgVal-----GlyLeuProGlyAlaGln 262  
 DB 361 GCCTTTTACACAGAACAGTTTTGAAGTAAAGAAATCTGCCAAGAGGT---CCTTCAGGTGAG 417  
 QY 263 GluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer----- 280

Db 418 GAGATTGTTTGCACCACTTATAATACTGGAACCGTGGAAATTCAGTGGTCAAGAGGGAAA 477  
QY 281 -----AsnAspGluLeuPheGlnThrPheCysAspPheProGlu 293  
Db 478 CATAAGGAAGTGGAGACCTGACAGAACAGAGCGTACAGTTATTTGTTGATTTCCCTGGAT 537  
QY 294 IleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeu 313  
Db 538 ATTATTGATGTCAGTATTAGCAAGCAACACGAG---GAATGCTCAATGAAGTAGAATT 594  
QY 314 ValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlu 333  
Db 595 GTAACCTGTCATTAACCAAGATGTTAAAGATTTTGGAGATAGAACCTTAGCTCATTTAAAGAA 654  
QY 334 AlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHis 353  
Db 655 GCCTTGTCATTCGTGTCATTAATGACGGGTATTACAGACTTAACCTGCGGATGGCCACAT 714  
QY 354 TyrPheCysLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHis 373  
Db 715 TACCTCTGCAAGAGGTGGCTCCCCAGCTGTGCTCGAGAACATACACAGCAACTGCCAC 774  
QY 374 GlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGly 393  
Db 775 GGCCCAATATCAATGGATTTTGCATTTAGCAAACTAAAGAGGGCGGTAAACCGACTGGA 834  
QY 394 ThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysVal 413  
Db 835 CTATATGTGTACATCGACCCCTAAGGACTTTCACAAATACCTTCTGACCTTGTCTGTT 894  
QY 414 GlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAla 433  
Db 895 GAGCGAGAAATGTCATTGAATAATAACACTGTTTGTATTACGAAGATGAGAATGGAGAA 954  
QY 434 PheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCys 453  
Db 955 TACAACCTCGCGGACTTAATGAAGCTTCAGTAACCTTAAGGACCTTTTGAATTTGTCTAC 1014  
QY 454 TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuThrSerCysCysAlaPro 473  
Db 1015 CAGATGGAACTGTGCGCTCGACAGCATCATCTTCAGTTTCCAAATGCTGCCCCCA 1074  
QY 474 ArgProLysGlyLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAla 493  
Db 1075 AAGCCAAAGATAAATCAAACTTCTCGTCTTCAGA-----ACAAATGGTATTCTGTAT 1128  
QY 494 ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr 510  
Db 1129 GTTCAGATCTCACCAACATTACAGAGCATATAAATGTGAATCAATGGTGTTCACAAA 1188  
QY 511 IleProThrAspSerLeuGluTyrPheHisGluAsnLeuGlyHisGlySerPheThrLysIle 530  
Db 1189 ATCAGGAATGAAGTTTAATATTAAATGAAGTCTTGCCCAAGTACTTTTACAAAATTT 1248  
QY 531 PheArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeu 549  
Db 1249 TTTAAAGGTGTAAAGAGAGAGTGGAGATTATGCTCACTGCACAAAAACCGAAGTCTTT 1308  
QY 550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSer 569  
Db 1309 TTGAAGTCTCTAGATAAAGCACATAGGAACCTATTTCAGAGTCTTTTTCGAAGCAGCAAGC 1368  
QY 570 LeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589  
Db 1369 ATGATGAGTCAGCTTTCTCAACAGCATTTGGTTTGAATTAATGTTGTTCTGTGCTGGA 1428  
QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608  
Db 1429 GAGGAGAACATTTCTGGTTCAAGATTTGTAATAATTTTGGATCACTGGATACATACCTGAAG 1488  
QY 609 LysArgGlyHisLeuValSerAlaSerTyrLysLeuGlnValThrLysGlnLeuAlaTyr 628  
Db 1489 AAGAACAAAAATTCATAAATATATTATGGAACCTTGGAGTGGCTAAGCAGTGGCATGG 1548

QY 629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648  
Db 1549 GCCATGCATTTCTAGAGAAATCCCTTATTTCATGGGAATGTGTGCTGCTAAAAATATC 1608  
QY 649 LeuLeuAlaArgGluGlyGly-----AspGlyAsnProProPheIleLysLeuSerAsp 666  
Db 1609 CTGCTTATCAGAGAAGAGACAGAGAACCGGGAACCCACTTTCATCAAACTTAGTGAT 1668  
QY 667 ProGlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpVal 686  
Db 1669 CCTGGCATATGACATTACAGTTCTACCGAAGGACATCTTCTCAGGAGAAATACCATGGGTA 1728  
QY 687 AlaProGlyCysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTyrGlyPhe 706  
Db 1729 CCTCTGAATGCATTGAGAAATCTAAATCTCGCAACACAGCAAGTGGAGCTTC 1788  
QY 707 GlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluPro 726  
Db 1789 GGGACCACTCTGTGGAGATCTGCAGTGGAGAGATAAGCCCTGAGTGTCTGTGATCT 1848  
QY 727 AlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTyrThrGlu 746  
Db 1849 CAAGAAGCTGAGTTCTATGAAGATAAGCATCAGCTTCTGCACCAAGTGGACAGAG 1908  
QY 747 LeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArg 766  
Db 1909 TTAGCAACCTTATAAATAATTCATGGACTATGAGCCAGATTTTCAGGCTCTGCTTCAGA 1968  
QY 767 AlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuLeuSerAspPro 786  
Db 1969 GCTGTATCGCTGATCTTAAACAGCTGTTTACTCCAGATTATGAACCTACTAACAGAA-- 2025  
QY 787 ThrProGlyIleProSerProArg---AspGluLeuCysValAlaGlyAlaGlnLeuTyr 805  
Db 2026 AATGACATGCTACCAACATAGAGAAATAGTGGCTTAGGTTTCTGTGTCT-----TTT 2079  
QY 806 AlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGly 825  
Db 2080 GAAGACAGGAGCCCTACACAGTTTGAAGAGAGACACTTGAAGTTTCTACAGCAGCTTGGC 2139  
QY 826 LysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGly 845  
Db 2140 AAAAGTAACTTCGGAGTGTGGAGATGCGCTGATGACCGCTGCAGGACAAACACTGGC 2199  
QY 846 ProLeuValAlaValLysGlnLeuGlnHisSerValProAspGlnGlnArgAspPheGln 865  
Db 2200 GAGGTGCTGCTGTGAAGAACTCCAGCAGCAGCTGAAGAGCACCTCCGAGACTTTGAG 2259  
QY 866 ArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyVal 885  
Db 2260 AGGAGATCAGATCCTGAATCCTTGCAGCATGACAACTCGTCAAGTACAGGAGTG 2319  
QY 886 SerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCys 905  
Db 2320 TGCTACAGTTCGGGTGCGCAACCTAAAGATTAAATATGGAATATTTACCATATGGAAGT 2379  
QY 906 LeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAspArgLeuLeuPhe 924  
Db 2380 TTACGAGACTATCTCCAAAAACATAAAGAACCGATAGATCACAAAAAATCTTCTCAATAC 2439  
QY 925 AlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAsp 944  
Db 2440 ACATCTCAGATATGCAAGGCGCATGGAATATCTTGTGTACAAAAAGGTATATCCACAGGAC 2499  
QY 945 LeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGly 964  
Db 2500 CTGGCAACAAGAACATATTTGGTGGAAATGAGAACAGGGTTAAATATAGGAGACTTCGGA 2559  
QY 965 LeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSer 984  
Db 2560 TTAACCAAGTCTTGCGCGAGGACAAAGATATCTCAAAAGTAAGGAGCCAGGGAAGC 2619

QY 985 ProfilePheTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAsp 1004  
 DB 2620 CCCATATTCGTGACGCACTCAATCCTTGACGGAGACCAAGTTTCTGTGGCCTCAGAT 2679  
 QY 1005 ValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSer 1024  
 DB 2680 GTGTGGAGCTTTGGAGTGGTCTTATACGAACCTTTTTCACATACATCGAAGAGTAAAGT 2739  
 QY 1025 ProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCys--- 1043  
 DB 2740 CCACCCGTGGAAATTATCGGAATGATTGGCAATGATAACCAAGGCAATGATTGTGTC 2799  
 QY 1044 ArgLeuLeuGluLeuAlaGluGlyArgArgLeuProProProThrCysProThr 1063  
 DB 2800 CATTTGATAGACTACTCAAGACCAACGGAAGATTGCCAAGGCCAAGAGTATGCCAGAT 2859  
 QY 1064 GluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPhe 1083  
 DB 2860 GAGATTTATGTGATCATCAGACAGTGTCTGGAACAACAATGTGAGCCGCTCCCTCCTTC 2919  
 QY 1084 AlaThrLeuSer 1087  
 DB 2920 AGGACCTTTTCG 2931  
 RESULT 15  
 ID AAD22680  
 XX AAD22680 standard; cDNA; 3495 BP.  
 AC AAD22680;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE Murine JAK2 tyrosine kinase cDNA.  
 XX  
 KW Murine; Src homology 2-Bbeta; SH2-Bbeta; neuroprotective; gene therapy;  
 KW cell differentiation; nerve regeneration; angiogenesis; embryogenesis;  
 KW cytosolic; antitense therapy; drug screening; cellular expression;  
 KW immunological disease; neurological disease; apoptosis; diabetes; cancer;  
 KW arthritis; JAK2 tyrosine kinase; ss.  
 XX  
 OS Mus sp.  
 XX  
 FH Key  
 FT CDS  
 FT Location/Qualifiers  
 FT 1..2982  
 FT /\*tag= a  
 FT /product= "Murine JAK2 tyrosine kinase protein"  
 XX  
 PN USG312941-B1.  
 XX  
 PD 06-NOV-2001.  
 XX  
 PF 26-NOV-1997; 97US-0980080.  
 XX  
 PR 26-NOV-1997; 97US-0980080.  
 XX  
 PA (UNMI ) UNIV MICHIGAN.  
 XX  
 PI Carter-Su C, Rui L, Karow DS;  
 XX  
 XX WPI; 2002-024907/03.  
 DR P-PSDB; AAB13587.  
 DR  
 XX  
 PT New isolated nucleic acid molecule encoding full length src homology 2  
 PT (SH2)-Bbeta protein, useful for modulating cellular expression of  
 PT active SH2-Bbeta protein -  
 XX  
 PS Example 1; Fig 3A; 71pp; English.  
 XX  
 CC The invention relates to human Src homology 2 (SH2)-Bbeta protein and its  
 CC DNA molecule. SH2-Bbeta is useful in drug screening assays designed to  
 CC identify drugs that interfere with the specific binding of JAK2 tyrosine  
 CC kinase which is a member of Janus family of tyrosine kinase-interacting  
 CC signalling molecule. SH2-Bbeta protein is useful for modulating,

CC preferably reducing cellular expression or intracellular concentration or  
 CC availability of active SH2-Bbeta. SH2-Bbeta is useful for treating a  
 CC disease associated with undesirable cell growth, differentiation, and  
 CC growth factor/cytokine, preferably interleukin, more preferably growth  
 CC hormone, platelet derived growth factor, nerve growth factor, epidermal  
 CC growth factor responsiveness, and for treating diseases associated with  
 CC cell movement. SH2-Bbeta molecule has important implications in cancer  
 CC metastasis, nerve regeneration, angiogenesis and embryogenesis and is  
 CC useful for preventing apoptosis, and treating diabetes, cancer,  
 CC arthritis, immunological diseases, neurological diseases, etc.  
 CC The invention also relates to compositions and methods for identifying  
 CC cytokine, hormone and growth factor signalling pathway agonists and  
 CC antagonists. Human SH2-Bbeta DNA is useful in detecting the presence of  
 CC SH2-Bbeta genes in gene transcripts, for detecting or amplifying DNA's  
 CC with substantial sequence similarity with SH2-Bbeta homologues and  
 CC structural analogues and for gene therapy applications. The present  
 CC sequence is murine JAK2 tyrosine kinase cDNA related to the invention.  
 XX  
 SQ Sequence 3495 BP; 1121 A; 688 C; 794 G; 892 T; 0 other;  
 Alignment Scores:  
 Score: 1.24e-170 Length: 3495  
 Pred. No.: 2360.00 Matches: 489  
 Percent Similarity: 65.89% Conservative: 164  
 Best Local Similarity: 49.34% Mismatches: 300  
 Query Match: 40.27% Indels: 38  
 DB: 24 Gaps: 15  
 US-09-397-967a-16 (1-1099) x AAD22680 (1-3495)  
 QY 128 IleLeuAspLeuHisValLeuGluHisPheAlaGlnHisArgSerAspLeuValSer 147  
 DB 1 CTGCTTGATGACTTGTGTCATGCTTACCTTCCCTCAGTGGCGCATGTTTGTTCAC 60  
 QY 148 GlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaVal 167  
 DB 61 GGATGATAAAAGTACCTGTGACTCATGAACTCAGGAAGAGTGTCTTGGGATGCGGTG 120  
 QY 168 LeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgProGlyGluLeuLeuLysThr 187  
 DB 121 TTAGACATGATGAGTAATAGCTAAGGAGAAAGACCACTCCACCTGGCTGTCTTAACTCT 180  
 QY 188 ValSerTyrLysAlaCysLeuProProSerLeuArgAspValIleGlnGlyGlnAsnPhe 207  
 DB 181 GTCAGTACAGACATCTTACCAAGTCGTTTCGCGCAAGATCCAGACATATCATT 240  
 QY 208 ValThrArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeu 227  
 DB 241 TTAACCCGGAAGCGAATCAGGTACAGATTTCGACAGATTCAATTCAGCAATTCAGTCAATGT 300  
 QY 228 ProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisPro 247  
 DB 301 AAAGCCACTGCCAGAACCTAAACCTTAAGTATCTTATAAACCTCGAACCCCTGCAGTCT 360  
 QY 248 AlaAlaThrThrGluThrPheArgVal-----GlyLeuProGlyAlaGln 262  
 DB 361 GCCTTCTACACAGACAGATTGGAAGTAAAGAAATCTGCAAGAGT---CCTTCAGGTGAG 417  
 QY 263 GluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSer----- 280  
 DB 418 GAGATTTTGGCAACCATTAATAACTGGAACCGGTGGAATTCAGTGTGTCGAAGGGAAA 477  
 QY 281 -----AsnAspGluLeuPheGlnThrPheCysAspPheProGlu 293  
 DB 478 CATAGGAAGTGAACACTGACACAGACAGACCGTACAGTTATATTGTGATTTCCCTCAT 537  
 QY 294 IleValAspValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeu 313  
 DB 538 ATTATGATGATGATTAAGCAAGCAATCAG---GAATGCTCACTCAAGTAAAGTAGATT 594  
 QY 314 ValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlu 333  
 DB 595 GTGACCGTCCACAGCAGGACGGGAAGGTCTTTGAAATAGAACTTAGCTCTTAAAGAA 654



```
Db      2800 CAAAGGCAATGATTGTGTTCATTGTAGAGCTACTGAAGAGCAACGGAAGATTGCCA 2859
Qy      1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076
Db      2860 AGGCCAGAAAGGATGCCCAAGATGAGATTATGTGATCATGACAGAGTGCTGTGGAACAACAAT 2919
Qy      1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087
Db      2920 GTGAGCCAGCGTCCCTCCTTCAGGGACCTTCC 2952
```

Search completed: February 4, 2004, 03:08:17  
Job time : 917.095 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 02:45:11 ; Search time 6446.54 Seconds  
(without alignments)  
4143.406 Million cell updates/sec

Title: US-09-397-967A-16  
Perfect score: 5860  
Sequence: 1 MAPPEETPLIPQRSCLSS.....RPAFATLSQLDPLWRGPG 1099

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09397967/runat\_03022004\_175632\_28586/app\_query.fasta\_1.1486  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967@cgn\_1\_1\_4514@runat\_03022004\_175632\_28586 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOC=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5669.5	96.7	3742	11	AK038268	AK038268 Mus muscu
2	5650	96.4	3781	11	AK088365	AK088365 Mus muscu
3	1447	24.7	4462	11	BC053021	BC053021 Mus muscu
4	1380	23.5	971	12	BT411962	BT411962 602966375
5	1358	23.2	1625	11	BC027234	BC027234 Mus muscu
6	1214.5	20.7	1200	13	BX415899	BX415899 BX415899
7	1192	20.3	826	12	B1557690	B1557690 603237050
8	1170.5	20.0	2222	11	AK087255	AK087255 Mus muscu
9	1141	19.5	698	12	BM935182	BM935182 UI-M-BH3-
10	1134	19.4	675	10	BF150250	BF150250 uy83e09.y
11	1081	18.4	677	14	BY730320	BY730320 BY730320
12	1071.5	18.3	927	13	BX370312	BX370312 BX370312
13	1032	17.6	655	14	BY746452	BY746452 BY746452
14	1031.5	17.6	907	12	BG870513	BG870513 602791432
15	1028.5	17.6	720	12	BG873355	BG873355 602794328
16	1016	17.3	621	9	AW742526	AW742526 up58a12.y
17	1012	17.3	731	9	AA920299	AA920299 vx93902.y
18	1007	17.2	666	14	BY734043	BY734043 BY734043
19	960	16.4	583	10	BB631706	BB631706 BB631706
20	935.5	16.0	871	13	B0709825	B0709825 AGENCOURT
21	930	15.9	559	12	BG800853	BG800853 0053-91 M
22	906	15.5	769	9	AJ442008	AJ442008 AJ442008
23	879	15.0	545	10	B8916148	B8916148 601666763
24	874	14.9	539	10	B8913247	B8913247 601668246
25	835.5	14.3	543	10	BF753801	BF753801 206963 MA
26	831.5	14.2	562	10	BF193178	BF193178 244670 MA
27	828	14.1	580	10	B8608771	B8608771 BB608771
28	824	14.1	557	12	B1339676	B1339676 364724 MA
29	823	14.0	544	10	BF705690	BF705690 243606 MA
30	807	13.8	590	14	W48204	W48204 mc87h11.r1
31	782.5	13.4	867	9	AA755769	AA755769 vv35a11.r
32	781.5	13.3	554	10	BF191625	BF191625 239261 MA
33	762	13.0	767	12	BT734179	BT734179 603351465
34	761	13.0	763	9	AJ399851	AJ399851 AJ399851
35	759.5	13.0	564	12	BM088287	BM088287 501721 MA
36	725.5	12.4	540	9	AW505513	AW505513 UI-HF-BN0
37	723	12.3	925	13	BQ683553	BQ683553 AGENCOURT
38	721	12.3	1097	12	BM455276	BM455276 AGENCOURT
39	720.5	12.3	526	10	BF191620	BF191620 239255 MA
40	709.5	12.1	593	9	AA881653	AA881653 vx21c02.r
41	705	12.0	501	10	BB623895	BB623895 ut14d09.y
42	704	12.0	696	14	CB246966	CB246966 UI-M-FIO-
43	702	12.0	436	9	AW141115	AW141115 EST291147
44	698	11.9	453	9	AA023709	AA023709 mh77e02.r
45	695.5	11.9	462	9	AW408832	AW408832 UI-HF-BM0

ALIGNMENTS

RESULT 1  
AK038268  
LOCUS  
DEFINITION Mus musculus 16 days neonate thymus cDNA, RIKEN full-length  
enriched library, clone:Al30091E14 product:Janus kinase 3, full  
insert sequence.  
ACCESSION AK038268  
VERSION AK038268.1 GI:26332476  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.





QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
DB 82 TCAGAGGAGGAGCCCTGCATGTGTCTCTCTCCCGGGACCTGGGCTCCCCAGCGA 141  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
DB 142 TTGTCAATCTCTTTTGGGACTACTTGGCTGAGGATTATGTGTGCGAGCTGCCAAGGCC 201  
QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
DB 202 TGTGGCATCTCCCTGTTTATCATTCGCTTTTGTCTCGGCCACTGAGACTTCTCTTCG 261  
QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
DB 262 TGGTTTCCCCCAAGCCACATCTTCTGCATAGAGGAGTGGACACTCAAGTCTTGGTCTAC 321  
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
DB 322 AGGCTACGCTTTATTTCCCTGACTGGTTTGGGCTGGAGACATGTCCCGCTTTGGGCTG 381  
QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
DB 382 CGCAAGAATTTGACCAGTCCCATCTTGACTTACAGCTTTTAGAACATCTCTTTGCTCAG 441  
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
DB 442 CACCCAGTGCCTGGTGGGCGCTCCCGGTGGGCTTAGCATGAAGAGCAGGGA 501  
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
DB 502 GAGTTCTCGAGCCTGGCCGTGTGGACTTGGCCCGACATGGCTCGTGAGCAGGCCAGCGC 561  
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
DB 562 CCAGGAGAGCTCTCTGAAGACGCTCAGTTACAAGCCTGTCTCGGCCCGAGCCTGGCGAT 621  
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
DB 622 GTGATCCAGGGCCAGAACTTCGTGACACGACGAGGCGCATCCGACGACCCGTGTGGCG 681  
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240  
DB 682 CTGCG-CCGTGTGTCTGCGCCAGCGCCGACCGCTACGCGCTCATGGCCAAATATATTTT 740  
QY 240 uAspLeuGluArgLeuHisProAlaThrThrGluThrPheArgValGlyLeuProGln 260  
DB 741 GGACCTGGAGCGGCTACATCCAGCGGCCACCCAGGAGACTTCCGFGTGGGGCTCCCGGG 800  
QY 260 yAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrpSerSe 280  
DB 801 CCCCCAGGAGGAGCCGGGCTTCTGGTGTGGCGGGGGAACCGGCATCTCTGGAGCTC 860  
QY 280 rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIl 299  
DB 861 CGGGGACCAAGGAGCTTTCCAGACCTTCTGTGACTTTCGGAAATCGTGGATGTACGAT 920  
QY 299 eAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMe 319  
DB 921 CAAGCAGGCCCCACGFTGGGTCCGACGAGGGGAACACCGGCTGGTCACCGTCAACAGGAT 980  
QY 319 tAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAl 339  
DB 981 GAGCGCCACATCTTGAAGCGGAGTTTCCGGGGCTGCCGTGAGGCGCTGTCTCTGTGGC 1040  
QY 339 aLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVa 359  
DB 1041 CCTTGTGGATGGTACTTCCGCTGATCTGCGACTCCAGGCAATATTTCTGCAAGGAGGT 1100  
QY 359 lAlaProProArgLeuLeuGluAlaAspValCysHisGlyProIleThrLeuAs 379  
DB 1101 GCGCGCCCGCGCTCTCGAGGAGGAGCGGAGCTGTGCCATGGACCCATCACGTTAGA 1160  
QY 379 pPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgAr 399

DB 1161 CTTTGCATCCACAGCTGAAGCGCGCTGGCTCCCTCCCGAGCACCTACATTCCTCCGCG 1220  
QY 399 gSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPr 419  
DB 1221 CAGCCCGCAGGACTATGACAGCTTTCTTTACCGCTGCGTCCAGACTCTCTTTGGCC 1280  
QY 419 oAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLe 439  
DB 1281 CGACTACAGGGCTGCCCTCATCCGCCAGACCCCGAGGGCTTTCTCCCTGGTTGGCT 1340  
QY 439 uSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuAr 459  
DB 1341 CAGCCAGCCCCACAGAGAGCTCGGGAGCTGTTGCAGCGCTCTGGAATTCCTGGGCTGCG 1400  
QY 459 gValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSe 479  
DB 1401 AGTAGCGTGTGCCCTGAACCTAACATCTCTGCTGCGCTGCCAGACCCCAAGAAAGTC 1460  
QY 479 rAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSe 499  
DB 1461 CAATTTGATCGTGGTGGAGGGGCTGCACCCCGCGCTGCCCTGGCTGCTCCCCGTC 1520  
QY 499 rCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHi 519  
DB 1521 CTGCTGTGCGCTGACACAGCTGAGCTTCCACACAATTCACACGAGCCTCGAGTGGA 1580  
QY 519 sGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVa 539  
DB 1581 CGAGAACCTGGGTACCGGTTCTTTTACCAAGATCTTCGTGGCCGCGAGCGGAGTGT 1640  
QY 539 lAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAs 559  
DB 1641 GGATGTTGAGACACATGACTCGAAGTCTCTCTGAAGGTATGGACTTCAGACATCGGA 1700  
QY 559 nCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLe 579  
DB 1701 CTGCATGGAGTCTTTCTGGAAGCCGCAAGCTTGATAGSCCAAGTATCTCTCCCGCACCT 1760  
QY 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599  
DB 1761 GGTGTTACTGCACGGCGTCTGCATGGCTGGAGACAGCATCATGTCGACAGAAATTTGTGTA 1820  
QY 599 rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLy 619  
DB 1821 TCTAGAGCAATGACATGACTCTGCGAAGCGTGGCCACCTGGTGTACGACAGCTGGAA 1880  
QY 619 sLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPr 639  
DB 1881 ACTGCAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTACTTGGAGGACAAAGGCCCTCC 1940  
QY 639 oHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGlyAsnPr 659  
DB 1941 TCACGGCAACGCTCTCAGCACGGAAGTGTCTCTGGCTGCTGAGGGGGGTGATGGAAATCC 2000  
QY 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetIe 679  
DB 2001 ACCTTTTCAATAGCTGAGTGATCTCGTGGTGTAGTCCCCTGCTGGTGTGGAGCCTGGAATGTCT 2060  
QY 679 uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699  
DB 2061 CACCCAGCAATACCTCGGTGGCCCCCGAATGTCTCCAGGAGGCTCAGACACTCGGCTT 2120  
QY 699 uGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl 719  
DB 2121 GGAGGCTACAAGTGGGGCTTTGGAGCCACACGCTGGGAGGTGTTTCAGCGGGGACCCGCG 2180  
QY 719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnIe 739  
DB 2181 CCACATCACTCGCTGGAGCGCCGCAAAAGCTGAAGTTCTATGAGGACCCAGGACAGCT 2240  
QY 739 uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr 759

Db 2241 GCCCGCTCTCAAATGGACAGAACTGGCGGAGCTTATCACAGTGCATGCGCTATGATCC 2300

Qy 759 oGlyArgArgProSerPheArgAlaLeuLeuArgAspLeuAsnGlyLeuLeuThrSerAs 779

Db 2301 TGGCGGCGCCCTCTTCGAGCTATCCCTCAGAGACTCAACGGCTCTATTACATCAGA 2360

Qy 779 pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVa 799

Db 2361 TTACGAGCTCTCTCAGACCCACACCTGGCATCCGAGTCCCTCGAGATGAGCTGTGC-- 2418

Qy 799 lAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLy 819

Db 2419 -GGTGGCGCCAGCTCTATGCTGCCAGGACCCGCCATATTCAGGAGAGACACCTTAA 2477

Qy 819 sTyrIleSerLeuLeuGlyIleGlyAsnPheGlySerValGluLeuCysArgTyrAspPr 839

Db 2478 GTACATCTCTTGTGGCGAAGGCACTTTGGCAGCGTGGAGCTGTGGCTATGACCC 2537

Qy 839 oLeuGlyAspAsnThrGlyProLeuValAlaValIleGlnLeuHisSerValProAs 859

Db 2538 CCTGGGGGACATACGGGACCCCTGTGGCAGTGTAACAGCTACAGCAGCGGGCCAGA 2597

Qy 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLeuAlaLeuHisSerAspPheI 879

Db 2598 CCAGCAGAGGAGCTTCCAGCGGAGATTCAGATCCTTAAGGCTCTGCACAGCGACTTCAT 2657

Qy 879 eValIleTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetG 899

Db 2658 CGTCAAGTACCGGGGAGTCACTATGGCCAGCGTGGCAGAGCTGGGTGGATGGA 2717

Qy 899 uTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly---LeuHisTh 918

Db 2718 GTACCTGCCAGCGCTCGCTCGCAGACTTCTCGCAGCCCATCGCGCGCTGCACAC 2777

Qy 918 rAspArgLeuLeuLeuPheAlaTrpGlnIleCysGlyGlyMetGluTyrLeuGlyAlaAr 938

Db 2778 CGACCGCTACTGCTGTTCGCTGGCAGATTCGACAGGCGATGAGTACCTGGTGGCGG 2837

Qy 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleuValGluSerGluAlaHisVa 958

Db 2838 CCGCTGGTACACGTCACCTGGCTGGCGGCAATCTTGTGGAGAGCGAGGCTCATGT 2897

Qy 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa 978

Db 2898 GAAGATCGCGACTTCGGCTCGCTAGCTGCTGCCCTCGGAAGGACTACTACGTGT 2957

Qy 978 lArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnI 998

Db 2958 CCGGAGCCTGGCCAAAGCCCATCTTTTGGTATGCCCGGAGTCCCTATCTGACACAT 3017

Qy 998 ePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTy 1018

Db 3018 CTTCTCCCGCAATCTCAGCTGTGGAGTTCGGAGTGGTGTGTACGAGCTCTTACCTA 3077

Qy 1018 rCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGl 1038

Db 3078 CTGCGACAGAGCTGCACCCATCCGCTAGTTCCTGAGCATGATGGGCTCGAGCTGA 3137

Qy 1038 uGlyProProLeuCysArgLeuLeuGluLeuAlaGluArgArgLeuProProPr 1058

Db 3138 AGGACCCCGCTCTGCCGCTCTCGGAGCTGCTGGAGGCGGCGAGCTCCACACACC 3197

Qy 1058 oProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHi 1078

Db 3198 TCCCACTGCCCCACGAGGTTCAGAGCTCATCAGCTGTGTGGGCGGCCACCCGCGCA 3257

Qy 1078 sAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrpArgGlyArgPr 1098

Db 3258 CGACCGGCCAGCCTTGGCCACCTTGAGCCCCCAGCTGGACGCGCTGTGGGTGGAAGACC 3317

Qy 1098 oGly 1099

Db 3318 CGGA 3321

RESULT 2  
AK088365 3781 bp mRNA linear HTC 05-DEC-2002  
LOCUS Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN  
DEFINITION full-length enriched library, clone:E430014A10 product:Janus kinase  
3, full insert sequence.

ACCESSION AK088365  
VERSION 1 GI:26353349  
KEYWORDS HIT; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636

REFERENCE 2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159

REFERENCE 3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11078661

REFERENCE 4  
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hata, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Stauber, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P.,  
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,  
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohseki, S.  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

REFERENCE 5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3781)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ono,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akaihi,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muratsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan. (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
 please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://phantom.gsc.riken.go.jp/  
 URL: http://location.qualifiers

## FEATURES

source

```
1. .3781
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430014A10"
/db_xref="taxon:10090"
/clone="E430014A10"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
63. .3362
/notes="unnamed protein product: Janus kinase 3
(MGI:MGI_99928, GB|NM_010589, evidence: BLASTN, 98%,
match=3715)
putative"
/codon_start=1
/protein_id="BAC40305.1"
/db_xref="GI:26353350"
/translation="MAKPSSEETPLIPQRSCSLSSSEAGALHVLPPRGPPQRLSFS
FDYLAEDLCVRAAKGLIFVTHSLFALATEDSCWFPSPSHFICLEDVDTQVLYRL
RYFPDWFLETCRGLRDLTSAILDHLEHFAQHRSDLVSRFLVGLSMKEQG
EFLSLAVLQMAQREAOQPGELLTYSYKALPSPSRDVIQGNFVRRIRRTV
LALRRVACQADRLAMKYLILDLERLHPAATTETPRVGLPGAEPGLLRVADNGI
SNSSGQELFTQFCDFPEIVDYSIKOAPRVGAGHRLVTVMDGHILLEAPGLPE
ALSFLVADGFRLLICDSRHYCKEVAPEPRLEERELCHGITLDFALHKLKAGSL
PGTYILRRSPQIDYLLTACVTPDIPGDPYKGLIRCPDPSGAFSLVLSQPHRSUREL
LAAWNSGLVDGAALNLSCCAPRPEKSNLIVRRGCTPAPAGSPSCCALTLQSLF
HTPTDLSLWHEHNGHSFTKIPRRREVVDGETHDSVLLKVMDSHRNOMESFLE
AASLSQVSPHLLHGVCMAGDSIMVOEFVVLGAILDMLYLRKRHLVSAKMLQVTK
QLAYALNYLEDKGLPHGNYSARKVLLAREGGDGNPPFIKLSDPGVSPTVLSLEMTDR
IPWAPCECLQEAQTLGLEADKWGFATWVEVSGGPAHTISLEPAKLFYEDQGLP
ALKWTEAGLITQCMAYDGRPSFRALIRDLNGLITSYELSDPTPGIPRDELC
GQAQLVADQDPAIFEEHLYKISLLGKNFSGVELCRYPDLGTPLVAVKLOHSG
PQQRDFQREIQTILKALHSDFIYKRVGSYGPQSLRVLMEYLPSCGLRDLQHRRA
RLHTRDLRLFWAGIICKGMVYLGARRCVHRDLAARNLIVSEAHVLIADFLAKLPLG
KDYVVRPEGQSPFWAPESLSDNIFPSQSDVMSGVVLYELFTYCDKSCSPSBEFL
SMWGPREGFPFLCLLELLAEGRRUPFPPTCTPTEVQELMQLCWAPSPHDPAPFGTLSP
QLDALWRGRPG"
3764. .3769
/notes="putative"
```

## CDS

```
polya_site 3781
BASE COUNT 707 a 1193 c 1106 g 775 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 3781
Score: 5650.00 Matches: 1077
Percent Similarity: 97.91% Conservative: 2
Best Local Similarity: 97.73% Mismatches: 18
Query Match: 96.42% Indels: 6
DB: 11 Gaps: 4
US-09-397-967A-16 (1-1099) x AK088365 (1-3781)
Qy 1 MetAlaProProSerGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20
Db 63 ATGGCACCTCCAAAGTGGAGGAGACCTCTGATCCCTCAGCGCTCTTGCAGCCTCTCATCC 122
Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 123 TCAGAGCGAGGAGCCCTGTCATGTGCTTCTTCCCGGGGACCTGGCGCTCCCGACGA 182
Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 183 TTGTCATTCTCTTTTGGGACTACTTGGCTGAGGATTATGTGTGCGAGCTGCCAAGGCC 242
Qy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 243 TGTGGCATCTCGCTGCTTTTATCATCTGCTTTTCCCTCGGCCACTGAGGACTTCTCTTGC 302
Qy 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 303 TGGTTTCCCGCCAAAGCCACATCTTCTGCATAGAGAGCGTGACACTCAAGTCTTGGTCTAC 362
Qy 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 363 AGGCTACGCTTTTATTTCCCTGACTGTTGGCTGGAGACATGTCACTGCTTTGGGCTG 422
Qy 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 423 CGCAAAAGATTTGACCAGTGCATCTTGTACTTACACGTTTAAACATCTCTTTTGTCTAG 482
Qy 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 483 CACCGCAGTGACCTGTGTGAGTGCGGCCCTTCCCGGTGGGCTTACGATGAAGGAGCAGGGA 542
Qy 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 543 GAGTTCTCTGAGCCTGGCCGCTGTGGACTTGGCCAGATGGCTCTGTGAGCAGGCCAGCGC 602
Qy 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 603 CCAGGAGAGCTGCTGAAGACGGTCACTTACAAGCCGTCTGCGGCCAGCGCTGCGCGAT 662
Qy 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220
Db 663 GTGATCCAGGGCCAGAACTCTGTGACACAGCGCGCATCCGAGGACCGTGTGTTGGCG 722
Qy 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240
Db 723 CTGGC-CCGTGTGGTGCCTGTCCAGCGCCACCGCTACGCGCTCATGGCCCAAGTATATTT 781
Qy 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
Db 782 GGACCTGGAGCGCTACATCCAGCGCCACACCGAGACCTTCCGTGTGGGCTCCCGCG 841
Qy 260 yAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyrSerSe 280
Db 842 CCCCCAGGAGGAGCGGGGCTTCTGCGTGTGGCGGGGAGCAACCGGCATCTCTCGGAGCTC 901
Qy 280 rAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIl 999
```

polyA\_signal

Db 902 CGGGACAGGAGCTTTTCCAGACCTTCTGTGACTTTTCCGGAAATCGTGGATGTGACAT 961  
Qy 299 eAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMe 319  
Db 962 CAAGCAGGCCCCACGTGTGGGTCCGGCAGGGGAAACACCGGTGTGACCGGTACACAGAT 1021  
Qy 319 tAspGlyHisIleLeuGlnAlaGluPheProGlyLeuProGluAlaLeuSerPheValAl 339  
Db 1022 GGACGGCCACATCTCTGGAGCGGAGTTTCCGGGCTGCTCCAGCGGCTGTCTCTGTGGC 1081  
Qy 339 aLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluVa 359  
Db 1082 CTTTGTGATGGGTACTTCCGCTGATCTCGACTCCAGGCATTATTCTGCAAGGAGGT 1141  
Qy 359 lAlaProProArgLeuLeuGlnGluAlaAlaAspValCysHisGlyProIleThrLeuAs 379  
Db 1142 GGCGCCCGCGGTGTGGAGGAGGAGCGGAGCTGTGCCATCGACCCATCAGCTTAGA 1201  
Qy 379 pPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgAr 399  
Db 1202 CTTTGCATCCACAAGCTGAAGCGCGTGGCTTCTCCAGGACCTACATCTCCGCGC 1261  
Qy 399 gSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyPr 419  
Db 1262 CAGCCCGCAGGACTATGACAGCTTCTTCTTACCGCTGCTGCCAGACTCCTCTTGGGCC 1321  
Qy 419 oAspTyrLysGlyCysIleuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLe 439  
Db 1322 CGACTACAAGGGCTGCCTCATCTCCGCCAGACCCAGCGGGCTTCTCCCTGGTGGCT 1381  
Qy 439 uSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTyrAsnSerGlyLeuAr 459  
Db 1382 CAGCCAGCCCCACAGAAGCTTCGGGAGCTGCTGACGCC---TGGAAATTCTGGGCTGG 1438  
Qy 459 gValAspGlyAlaAlaLeuTyrLeuThrSerCysAlaProArgProLysGluLysSe 479  
Db 1439 AGTAGCGGTCTGCCCTGAACTTAACATCTGCTGCGCTCCAGACCCCAAGGAAAGTC 1498  
Qy 479 rAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSe 499  
Db 1499 CAATTGTATCTGTGTGGAAGGGCTGCACCCCGCGCTGCTCCCTGCTGCTCCCGCTC 1558  
Qy 499 rCysCysAlaAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrPi 519  
Db 1559 CTGCTGTGGCTGACACAGCTGAGCTTCCACACATTTCCACGGACAGCTTGGAGTGCA 1618  
Qy 519 sGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValVa 539  
Db 1619 CGAAACCTGGGTACCGGTCTCTTTTACCAGATCTTCCGTGGCCGCGGAGGTCGT 1678  
Qy 539 lAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAs 559  
Db 1679 GGATGGTGAGACATGATGACTCGGAAGTCTCTCTGAAGGTCTAGGACTCCAGACATCGGAA 1738  
Qy 559 nCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisIle 579  
Db 1739 CTGCAATGGAGCTCTTTCTGGAGCGCGAGCTTGTATGAGCCAGTATCTTACCCGCACT 1798  
Qy 579 uValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTy 599  
Db 1799 GGTGTACTGCACGCGCTGCTGCATGGCTGGAGACAGCATCATGTTGTCAGGAATTTGTGTA 1858  
Qy 599 rLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTyrPi 619  
Db 1859 TCTAGGACAAATGATGATGATCTCCGCAAGCGTGGCCACCTGTGTGTCAGCCAGCTGGAA 1918  
Qy 619 sLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPr 639  
Db 1919 ACTGCAGGTGACCAAGCAGCTGGCATATGCCCTTAATCTTGGAGGACAAAGGCTTCC 1978  
Qy 639 oHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyAspGlyAsnPr 659  
Db 1979 TCACGGCAACGTCTCAGCAGGAAGGTGCTCCTGCTGCTGAGGGGGGTGATGGGAATCC 2038

Qy 659 oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe 679  
Db 2039 ACCTTTTCAATTAAGCTGAGTGATCTGTGTGTCACTCCCACTGTGTGAGCCTGGAAATGCT 2098  
Qy 679 uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe 699  
Db 2099 CACCGACAGAAATACCTCTGGGTGGCCCGGAATGTCTCAGGAGGCTCAGACACTCGGCTT 2158  
Qy 699 uGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl 719  
Db 2159 GGAGCTGTGACAAGTGGGCTTTTGAGGCCACACATGGGAGGTGTTCAGCGGGGACCCGC 2218  
Qy 719 aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLe 739  
Db 2219 CCACATCACCTCTGCTGAGCCCGCCCAAAAGCTGAGTTCTATGAGGACCCAGGACAGCT 2278  
Qy 739 uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr 759  
Db 2279 GCCCGCTCTCAATGGACAGAACTGGCGGACTTATCACACAGTGCATGGCTATGATCC 2338  
Qy 759 oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs 779  
Db 2339 TGGCGGCGCGCTCTTCCAGCTATCTCAGAGACCTCAACGGCTCATTTACATCAGA 2398  
Qy 779 pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysVa 799  
Db 2399 TTACGAGCTCTCTCAGACCCCACTGGATCCCGAGTCTCCGAGATGAGCTGTGC-- 2456  
Qy 799 lAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLy 819  
Db 2457 -GGTGGCGCCAGCTCTATGCTGCCAGGACCCCGCCATATTCAGGAGAGACACCTTAA 2515  
Qy 819 sTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPr 839  
Db 2516 GTACATCTTCTGTGGGCAAGGCAACTTTGGCAGCGTGGAGTGTGCCCTATGACCC 2575  
Qy 839 oLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAs 859  
Db 2576 CCTGGGAGCAATACGGGACCCCTGTGTGGCACTGAACAGCTACAGCACAGCGGCCAGA 2635  
Qy 859 pGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIl 879  
Db 2636 CCAGCAGAGGACTTCCAGCGGAGATTACAGATCTTAAGCTCTGCACAGCGACTTCAT 2695  
Qy 879 eValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGl 899  
Db 2696 CGTCAAGTACCGGGAGTCACTATGGGCCAGGTCCGACAGGCTTCCGCTTGGTGTGATGA 2755  
Qy 899 uTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---LeuHisTh 918  
Db 2756 GTACTGCCAGCGCTGCTTCCGAGACTTCTGCAGCGCCATCGCGCGCTGCACAC 2815  
Qy 918 rAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaAr 938  
Db 2816 CGACCGCTACTGCTGTTCGCTTGGCAGATCTGCAAGGGCATGGAGTACCTGGGTGCGCG 2875  
Qy 938 gArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVa 958  
Db 2876 CCGTGGCTGACACCGTGCCTGGCTGCGCGCAACATCTTGTGGAGAGCGAGGCTCATGT 2935  
Qy 958 lLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVa 978  
Db 2936 GAAGATCGCGGACTTCGGCTCGTAACTGCTGCCCTCGGGAAAGGACTACTACGTGCT 2995  
Qy 978 lArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIl 998  
Db 2996 CCGCAGGCTGGCCAAAGCCCATCTTTTGTGATGCCCGGAGTCCCTATCTATCGAACAT 3055  
Qy 998 ePheSerArgGlnSerAspValTyrPheGlyValValLeuTyrGluLeuPheThrTy 1018  
Db 3056 CTTCTCCCGCAATCTGACGTGTGGAGCTTCGGAGTGTGTGTGTACGAGCTCTTCCACTA 3115



QY 180 ArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProSerLeuArg 199  
Db 783 CCCTGGAGGAGTGGCCAGAGATCAGCTTCAAGAACTGCATCCCTCATCTCCGA 842  
QY 200 AspValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeu 219  
Db 843 CAGCACATCCGGCAGCAATGTCTCACACGCTGGCGTCCACAGA--GTCTCCGC 899  
QY 220 AlaLeuLeuProCysGlyArgLeuProGlyArg-----ProTyrAlaLeuMetAlaLys 237  
Db 900 CGCTTCCTGGGGCTCCCG--CCYGGCCACTCTCCAGCAGGTGTGTGTGAAG 956  
QY 238 TyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGly 257  
Db 957 TACTTGGCTACCTGGAGCGCTGGCTCCGGC----- 989  
QY 258 LeuProGlyAlaGlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePro 277  
Db 990 -----TTTGGCTCAGAGCGCATACCT 1010  
QY 278 TrpSerSerAsnAspGluLeuPheGlnThrPheCysAspPheProGluIleValAspVal 297  
Db 1011 -----GTGTGTCATCTG 1022  
QY 298 SerIleAsn-GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValTh 317  
Db 1023 GAGGTCTGGCAGACCCGAGAGGACCCTGCTATACCCAGA-----ACAGTGGG 1073  
QY 317 rArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPh 337  
Db 1074 CAGACCGCTGGGGATCCAGGCCAGAGCTGCTTCCGGGCTCCACCCAGAGTACTG 1133  
QY 337 eValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLy 357  
Db 1134 GTGACAGGCACCGGAGGTATCCAGTGGCATCCGCTGCAGACCCAGGGATCCGTTTGTACA 1193  
QY 357 sGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleTh 377  
Db 1194 G-----GCCAAGCTGTGCCAGAGGAC----- 1215  
QY 377 rLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrThrIleLe 397  
Db 1216 -----GGCCTCTACCTGAT 1229  
QY 397 uArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrProLe 417  
Db 1230 TCAGTGGAGCACCGCCACCTGCACCGCTGTACCTTACCGTGGCCCATCGAAACCCCGC 1289  
QY 417 u-----GlyProAspTyrLysGlyCysLeuIleArgGlnAspPro----- 430  
Db 1290 TTTTCAGTAATGGCCCT-----CGGGGCTGGCGCTCGAAAGTTCCCCCATCACACAGA 1343  
QY 431 -SerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLe 450  
Db 1344 GCCTGGAGCCTTTGTCTGGATGGCTGGCGCGCTCTTTGCCAGCTTGGGGGACCTTCG 1403  
QY 450 uAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCy 470  
Db 1404 GCTCGCCTTGCAGGCGCTGCTGCTCGGGCCGGTGTGATGACTGTTTCCCTTCACACACTG 1463  
QY 470 sCysAlaProArgProLys----- 476  
Db 1464 CTGCTGCCCCCGCCAGAGGTTGATAGAGAGTGGGAAACACAGGGTAGGATGAGTCT 1523  
QY 477 -----GluLysSerAsnLe 481  
Db 1524 GTACCCCTGAAGCCACCCACAGGGCGATGTCCACTCCTGTGCTAGAAATCTCCAACT 1583  
QY 481 uIleValValArgArgGlyCysAsnProAlaProAla----- 493  
Db 1584 GTCTATATGCGGGGTCTAG--GGCCACACCCCGCTCTCAACCTCAGTCAGTCAGCT 1642  
QY 494 -----ProGly----- 495

Db 1643 TCCACAGGGTTCCACAGGATGAATACCCAGGGTGAAGTGTCTCTGGCAGGACCTGCAGA 1702  
QY 496 -----CysSerProSerCysCysAlaLeuThrGl 505  
Db 1703 CTCTGGGGGGGGGGGTGCACCTGTATCCAGTGCACCCACACTGC----- 1748  
QY 505 nLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGl 525  
Db 1749 -ATTCTGTACCACCTTCTTACAGCTG-----TCCCACTTTGGGCAAGG 1789  
QY 525 ySerPheThrIlysIlePheArgGlySerArgArg-----GluVa 538  
Db 1790 CACAAGACCAATGTGTATGAGCGCTTCTAAGAGTGGGAGGCCGCCGATGAGGGCAAGT 1849  
QY 538 lValAspGly-----GluThrHisAspSerGluValLeuLe 550  
Db 1850 GGACAATGGATGTCCCTCCTGAGCTGTGGGACTAGTGGGAGCAGCTTCAGTGGTGTCT 1909  
QY 550 uLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLe 570  
Db 1910 CAAGTTCTCGACCCCATGCAC----- 1931  
QY 570 uMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAs 590  
Db 1932 -----CATGATATCGC-CTT----- 1945  
QY 590 pSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysAr 610  
Db 1946 -GATATCATTTGTGACAGAAATTCGAGAACATGGTCCCTTGGATGTGTGTTACGGCGACA 2004  
QY 610 gGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLe 630  
Db 2005 GAGGGGCCAAGTGGCCCATGACCTGGAGATGGTGTGCTCAGCAGCTGGCCAGCGCTCT 2064  
QY 630 uAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLe 650  
Db 2065 CAGTACTGTGAGGACAAAGATCTGGTTCACGGAAATGTATGTGGCCGGAACATCTCTCT 2124  
QY 650 uAlaArgGluGly--GlyAspGlyAsnProPheIleLysLeuSerAspProGlyVa 669  
Db 2125 GGCTCGCTGGGGTGGAGGAGGTACCAACCCCTTCATCAAGTAACTGATCTCTGGTGT 2184  
QY 669 lSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProThrValAlaProGl 689  
Db 2185 GGGCCAGGGTCCCTCTCCAGGGAAGAGCGGGTGGAGCGCATCCCTCGACAGCTCCCGA 2244  
QY 689 uCysLeuGln--GluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAl 708  
Db 2245 GTGCTGTCTGGAGGGACCATAGCTTGGGTACTGCCACGGACATGTGGGGCTTTGGTGC 2304  
QY 708 aThrThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLy 728  
Db 2305 CACCTCTCTGAGATCTCTTTGATGGGAGCGACCCCTGCAGGGTCTGTCTCTCTGA 2364  
QY 728 sLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAl 748  
Db 2365 GAAAGAACGGTTCTACACAAAGAAACATCAGCTGCCTGAACCTCATCCCCAGAGCTGGC 2424  
QY 748 aGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaIl 768  
Db 2425 CACACTCAACCGCCAGTGCCTGATGAACACAGCAGCGGGCCATCTTCCGACCAT 2484  
QY 768 eLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuSerAspProThrPr 788  
Db 2485 TTTGGGAGCTTCACAGGCTG-----CAGCCAGAAATCTAGTCGCGACTTCGGCTGT 2538  
QY 788 oGlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGl 808  
Db 2539 GAACCTCAGACTCACCA-----GCATC 2559  
QY 808 nAspProAlaIlePheGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAs 828  
Db 828 -----GCATC 2559



```
Db 2560 AGACCCCACTGTTTTCACACGCGCTATTTGAAAAGATCCGGGATTTGGGTGAGGGTCA 2619
Qy 828 nPheGlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuVal 848
Db 2620 CTTTGGCAAGTACGCTGACTGCTACGACCCCAACCAATGACCGCACTGCGGAGATGGT 2679
Qy 848 lAlaValGlyGlnLeuGlnHisSerValProAspGlnGlnArgAsp---PheGlnArgG1 867
Db 2680 GGCCTGTGAAGCCCTGAAGGAAGGTCGGTCCCGAGCTCCGCTCAGCTGGCAGCGGA 2739
Qy 867 nIleGlnIleuLeuValAlaLeuHisSerAspPheIleValGlyTyrArgGlyValSerTy 887
Db 2740 GATCGAGATCTCGGACGCTGTACCAACGACATATTGTCAAGTATAAAGGCTCTGTGA 2799
Qy 887 rGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuAr 907
Db 2800 GGACCAAGGAGAGAGTCTGTACAGCTGGTCATGGAGTACGCTCCCTCTGGGACGCCCTCCG 2859
Qy 907 gAspLeuLeuGlnArgHis---ArgGlyLeuHisThrAspArgLeuLeuLeuPheAlaTr 926
Db 2860 AGACTACCTGCCAAGGCACTGCTAGGCTG-----GCGCAGCTCTGCTGTTGCCCA 2913
Qy 926 pGlnIleCysGlyGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAl 946
Db 2914 GCAGATCTGGGAGGCGATGCCCTACCTGCACGCTCAGCACTACATTCACCGAGACCTCGC 2973
Qy 946 aAlaArgAsnIleuValGluSerGluAlaHisValGlyIleAlaAspPheGlyLeuAl 966
Db 2974 CGCGCGCAAGCTGCTGTGCAACGACAGCTGGTCAAGATTGGAGACTTTGGCCTAGC 3033
Qy 966 aLysLeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerPro11 986
Db 3034 CAAGGCTGTACCTGAAGGCCACGAGTACTACCGAGTGGCGAGGACGGGACGCCAGT 3093
Qy 986 ePheTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTr 1006
Db 3094 GTTCTGTATGCCCCAGATGCTCCGAAGGAGTGCATAATTTTACTATGCACTGATGCTG 3153
Qy 1006 pSerPheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSe 1026
Db 3154 GTCTTGGGGTAACTCTGTATGAGTTGTGATACACTGTGACTCTAACCAAGAGTCCCCA 3213
Qy 1026 rAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLe 1046
Db 3214 TATGAATTCACCGAGTCTATCGGCATACCCAGGGCCAGATGACCGTGTGAGGCTCAC 3273
Qy 1046 uGluLeuLeuAlaGluGlyArgArgLeuProProProProProProProProProProPro 1066
Db 3274 AGAGCTGCTGGAACGAGGAGAGAGGCTGCTCGGCTGACAGGTTGCTCCCTGTGAGATCTA 3333
Qy 1066 nGluLeuMetGlnLeuCysTyrAlaProGluProHisAspArgProAlaPheAlaThrLe 1086
Db 3334 TCACCTCATGAAGAATCTGGGAGACAGAGGCTCTTCCGCGCCACCTTCCAGAACTCT 3393
Qy 1086 uSerProGlnLeu 1090
Db 3394 TGTGCCCATCTCTC 3406

RESULT 4
BI411962
LOCUS 602966375F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5121786 5',
DEFINITION mRNA sequence.
ACCESSION BI411962
VERSION BI411962.1 GI:15172885
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Eutelestomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 971)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
```

JOURNAL  
COMMENT

Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1297 row: e column: 19  
High quality sequence start: 25  
High quality sequence stop: 785.  
Location/Qualifiers

## FEATURES

source

1. .971  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6 J"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5121786"  
/tissue\_type="pooled lung tumors"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NCI\_CGAP\_Lu33"  
/notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st  
strand cDNA was prepared from mRNA obtained from pooled  
lung tumors with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGAGCGCGCTCTGTTTTTTTTTTTTTTT 3'].  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not  
I and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 216 a 274 c 279 g 202 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1,94e-111 Length: 971  
Score: 1380.00 Matches: 287  
Percent Similarity: 89.30% Conservative: 5  
Best Local Similarity: 87.77% Mismatches: 27  
Query Match: 23.55% Indels: 10  
DB: 12 Gaps: 2

US-09-397-967A-16 (1-1099) x BI411962 (1-971)

Qy 475 ProLysGluLysSerAsnLeuValValArgArgGlyCysAsnProAlaPro 494  
Db 6 CCAAGGTCAAAGTCATTTTGTGCTGGTGGAGGGGCTGCACCCCGCGCTGCCCT 65  
Qy 495 GlyCysSerProSerCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAsp 514  
Db 66 GGCTGCTCCCGCTCTGCTGTGCTGCACACAGCTGAGCTTCCACACAATTCACCGGAC 125  
Qy 515 SerLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrIlePheArgGlySer 534  
Db 126 AGCTTGAGTGGACGAGAACCTGGGTACGGTCTCTTTTACCAAGATCTTCCGTCGCGC 185  
Qy 535 ArgArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuValMetAsp 554  
Db 186 AGCGGGAGGTCGTGGATGTGAGACACATGACTCGAAGTCTCTCTGAAGGTTCATGGAC 245  
Qy 555 SerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnVal 574  
Db 246 TCCAGACATCGGAACCTGCATGGAGTCTTTCTGGAAGCGCGCAAGCTTGATGAGCCAA 305  
Qy 575 SerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal 594  
Db 306 TCTACCCCGACCTGGTGTACTGACCGCGCTGTGATGGCTGGACAGCATCATGGTG 365  
Qy 595 GlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuVal 614

366 CAGGAATTTGTGTATCTAGGAGCAATTTGACATGTACCTGGCCAGCGTGGCCACCTGGT 425  
 615 SerAlaSerTrpIysLeuGlnValThrIysGlnLeuAlaValAlaLeuAsnTrpLeuGlu 634  
 426 TCAGCCAGCTGGAAACTGCAGGTGACCAAGCAGCTGGCATATGCCCTTAACCTATTGGAG 485  
 635 AspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly 654  
 486 GACAAAGGCTTCTCTACCGCAAGCTCTACAGCAGGAAGTCTCTGGCTCTGGAGGGG 545  
 655 GlyAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeu 674  
 546 GGTGATGGGAATCCACCTTTTCAATTAAGCTGAGTGATCTGTGTGTCAGTCCCACTGTGCTG 605  
 675 SerLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAla 694  
 606 AGCTGGAAATGCTCACCGCAGAAATACCTGGGTGGTCCCGCAATGTCACGAGGCT 665  
 695 GlnThr-LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPh 714  
 666 CAGCAACTCTGCTTGGAGGCTGACAAGTGGGCTTTGGAGCCACCACTGCGGAAGTGT 725  
 714 eGlnArgGlyProAlaHisIleThrSerLeuGlu-ProAlaLysLysLeuLysPheTrpG 734  
 726 CAGCGGGGACCGCCCAATCACCTGCTGGAGCGCCGCCCAAAAGCTGAAGTCTCTATG 785  
 734 luAspGlnGlyGlnLeuProAlaLeu---LysTrpThrGluLeuAlaGly-LeuIleTh 752  
 786 AGGACCAAGGACCAAGTCCCGGCTCTCAATGGCAGCAAGTGGCCCGGAACCTTATCAC 845  
 752 rGlnCys-MetAlaTrpAspProGlyArgArgProSerPheArgAlaIleLeu-ArgAsp 771  
 846 ACAGTGCAATGGGTATGAATCCTGGCGGGCCCTTCATTCGAGATATCTCTCAGAGAC 905  
 772 LeuAsnGlyLeuIleThrSerAspTrpGluLeuLeuSerAspProThrProGlyIlePro 791  
 906 CTCACGCGCT-ATTACATTGGTTTACGAGCTGCTCTA-AAACCCCAAACTGGAT---CCC 960  
 792 SerProArg 794  
 961 AGTCGCGAGA 969

BC027234 1625 bp mRNA linear HTC 04-MAR-2003  
 Mus musculus, Similar to Janus kinase 3, clone IMAGE:3489805, mRNA.  
 BC027234  
 BC027234.1 GI:20071020  
 HTCC

SOURCE  
 Mus musculus (house mouse)

ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1625)  
 Strausberg, R.  
 Direct Submission  
 Submitted (04-APR-2002) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Baylor College of Medicine Human Genome  
 Sequencing Center  
 Center code: BCM-HGSC  
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
 Contact: [amg@bcm.tmc.edu](mailto:amg@bcm.tmc.edu)  
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleseg, H.,  
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,

A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.llnl.gov>  
 Series: IRAK Plate: 35 Row: g Column: 19  
 This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein  
 This clone has the following problem: retained intron.

## FEATURES

source  
 Location/Qualifiers  
 1..1625  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3489805"  
 /tissue\_type="Mammary tumor. WAP-TGF alpha model. 7 months old, gross tissue."  
 /clone\_lib="NCI\_CGAP Mam5"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"  
 328 a 470 c 483 g 344 t

BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,71e-109 Length: 1625  
 Score: 1358.00 Matches: 281  
 Percent Similarity: 75.59% Conservative: 7  
 Best Local Similarity: 73.75% Mismatches: 11  
 Query Match: 23.17% Indels: 84  
 DB: 11 Gaps: 3

US-09-397-967A-16 (1-1099) x BC027234 (1-1625)

QY 1 MetalProProSerGluThrProleuIleProGlnArgSerCysSerLeuSerSer 20  
 DB 42 ATGGCACCTCCAAAGTGAGGAGACACCTCTGATCCCTCAGCGCTCTTGGAGCTCTCATCC 101  
 QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProArgGlyProGlyProGlnArg 40  
 DB 102 TCAGAGGAGGAGGCGCTGCATGTGCTCTCTCCCGGGGACCTGGCGCTCCCGAGCGA 161  
 QY 41 LeuSerPheSerPheGlyAspTyrIleuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
 DB 162 TTGTCAATCTCTTTTGGGAGCTACTTGGCTGAGGATTTATGTGTGCGAGCTGCCAAGGCC 221  
 QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 DB 222 TGTGGCATCCTGCGCTTTATCAATTCGCTTTGGCTCTGGCCACTGAGACTTCTCTTGC 281  
 QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 DB 282 TGGTTTCCCGCCAGCCACATCTTCTGCATAGAGGAGCTGGACACTCAAGTCTTGTCTAC 341  
 QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
 DB 342 AGGCTACGCTTTATTTCCCTGACTGGTGTGGCTGGAGACATGTACCGCTTTGGGCTG 401  
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
 DB 402 CGCAAGAATTTGACCAGTGCCATCTTACATGTTTGTAGAACACAGTCTTGTCTCAG 461  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
 DB 462 CACCGCAGTGACTGTTGAGTGGGCGCTCCCGTGGGCTTTAGCATGAAGGAGCAGGGA 521  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
 DB 522 GAGTTCTCTGAGCCTGGCGCTGTGACTTGGCCAGATGCTGCTGCTGAGCAGCCCGC 581  
 QY 181 ProGlyGluLeuLeuLysThrVal----- 188  
 DB 582 CCAGGAGAGCTCTCTGAAGACGCT-CAGGTGAGGGGCTCGGGCCACCCCTTCGTGCTTTC 640



```

188 ----- 188
641 TTGGGCTCGGATCTAGGGCCCTATCGTTTCCCTTACTTGAACACATACCTCTTCCCA 700
188 ----- 188
701 ACTTGTGACGACCACTCTCCGTCAGCCTGCCCTCGCGAAAGGTCCCTTATGTGTCT 760
189 -----Ser 189
761 CGTGTCTCTTGGGGCCCTCACCCTTAAGAGTTCGGCGTGCCTCCCTCCCTCCACAGT 820
190 TTYLYAAlaCysLeuProSerLeuArgAspValIleGlnGlnGlnGlnGlnGlnGlnGln 209
821 TACAAAGCTGTCTGCGGCCACCTCGCGGATGTATTCAGGGCCAGAACTTCGTGACA 880
210 ArgArgArgIleArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuProGly 229
881 CGCAGGCGCATCCGCGAGACCGTGTCTTGGCGTGGC-CCGTGTGTCTGCTGCCAGGC 939
230 ArgPro-TyrAlaLeuMetAlaLysTyrIleLeuAspLeuGluArgLeuHisProAlaAl 249
940 CGACCGCTACGCGCTCATGCCAAGATATCTGGACCTGGAGCGGCTACATCCAGCGGC 999
249 aThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlnGluProGlyLeuLeuAr 269
1000 CACCACCGAGACCTTCCTCGTGTGGGCTCCGGGCGGCCAGGAGCGCGGGCTTCTGCG 1059
269 gValAlaGlyAspAsnGlyIleProTrpSerSerAsnAspGluLeuPheGlnThrPheCy 289
1060 TGTGGCGGGGACACAGCATCTCTCTGAGCTCCGGGACCAG----- 1102
289 saaspPheProGluIleValAspValSerIle---AenGlnAlaProArgValGlyProAl 308
1103 -----GAGGTCTGGGCTGGGCTTGAGAGCGGGCTCCCGGGAGCGGTGGC 1152
308 a 308
1153 G 1153

RESULT 6
BX415899 1200 bp mRNA linear EST 13-MAY-2003
LOCUS BX415899 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP008YM18
DEFINITION 5-PRIME mRNA sequence.
ACCESSION BX415899
VERSION BX415899.1 GI:30641720
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
I. (bases 1 to 1200)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 9047.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0CAP008YM18&cluster=9047.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0CAP008BG09Q1.
FEATURES
Location/Qualifiers
1..1200
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="CS0CAP008YM18"
/tissue_type="THYMUS"
/clone_lib="Homo sapiens THYMUS"
/Note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."
BASE COUNT 247 a 337 c 340 g 234 t 42 others
ORIGIN
Alignment Scores:
Pred. No.: 1..15e-96 Length: 1200
Score: 1214.50 Matches: 245
Percent Similarity: 83.59% Conservative: 25
Best Local Similarity: 75.85% Mismatches: 51
Query Match: 20.73% Indels: 6
DB: 13 Gaps: 1
US-09-397-967A-16 (1-1099) x BX415899 (1-1200)
QY 289 CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAla 308
Db 89 TGCAGCTTCCAGAAATCGTAGACATTAGCATCAAGCAGGCCCGCGCGTGTGGCCGGCC 148
QY 309 GlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhe 328
Db 149 GGAGACACCGCTGTGTCTGTACAGGACAGACCAACAGATTTTAGAGGCCGAGTTTC 208
QY 329 ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle 348
Db 209 CAGGGCTCCCGAGGCTCTGTCTGTGGGGCTCGTGACGGCTACTTCCGGGTGACC 268
QY 349 CysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu 368
Db 269 ACGGACTCCAGCACTTCTTCTGCAAGAGGTGGCACCCCGCAGGCTGTGGAGGAGTG 328
QY 369 AlaAspValCysHisGlyProIleThrLeuAsp-PheAlaIleHisLysLeuLysAlaAl 388
Db 329 GCCGAGCAGTGCACGGCCCATCTCTGGACTTKGSSRVKCAACAAAGCTCAAGACTGG 388
QY 388 aGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLe 408
Db 389 GGGCTACGCTCTGGCTCTCTATGTTCTCCGCGCAGCCCCCAGGACTTTTCAGCTTCT 448
QY 408 uLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuLeuArgGl 428
Db 449 CTTCACTGTCTGTCTCCAGAACCCCTTGTCTCTGATATATAGGGCTGCTATCCGGCG 508
QY 428 nAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArgGl 448
Db 509 CA--CCACAGGAACCTTCTTCTGTGTGGCTCAGCCAGCCACCCACAGCAGTCTTCGAGA 566
QY 448 uLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuTh 468
Db 567 GCTCTGGCAACCTGCTGGGATGGGGGCTGCACGTAGATGGGGTGGCAGTCACTCCAC 626
QY 468 rSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgGlyCy 488
Db 627 TTCCTGTGTATCCCGCAGACCCAAAGAAAGTCCAACTGATCTGTGTCCAGAGAGTCA 686
QY 488 sAsnProAlaProAlaProGlyCysSerPro---SerCysCysAlaLeuThrGlnLeuSe 507
Db 687 CAGCCACCCCATCATCTCTGTGTTTCAGCCCCCAATCCCAATACCACTGAGTCAGATGAC 746
QY 507 rPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHisGlySerPh 527
Db 747 ATTTTCAAGATCTCCCTGCTGTGACAGCTGTGAGTGGATGAGAACCTGGGCCATGGTCT 806
QY 527 eThrLysIlePheArgGlySerArgGluValValAspGlyGluThrHisAspSerGl 547
Db 807 CACCAGATTTTACCGGGGCTGTGCGCATGAGTGTGTGATGGGGAGGCCCGCAAGACAGA 866

```

```

QY 547 uValLeuLeuValMetaspSerArgHisArgAsnCysMetClnSerPheLeuGlnAl 567
Db 867 GGTGCTCTGAAGTCAATGATGCACAGCAAGAACTGATGAGTCATTCCTCGGAAGC 926

QY 567 aAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHisGlyValCysMe 587
Db 927 AGCGAGCTTATGAGCCCAAGTGTCTACCGGCATCTCTGCTGCTCCACGGCGTGTGAT 986

QY 587 tAlaGlyAspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLe 607
Db 987 GGCTGGAGA-MAGACCATGGK-GAGGAATTGTACA-CTGGGGCCCATARAMATGTATCT 1043

QY 607 uArgLys 609
Db 1044 GCGAAAA 1050

RESULT 7
BI557690 826 bp mRNA linear EST 05-SEP-2001
LOCUS 603237050F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5289949 5',
DEFINITION mRNA sequence.
ACCESSION BI557690
VERSION BI557690.1 GI:15445004
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 826)
NITH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM11733 row: d column: 14
High quality sequence stop: 760.
FEATURES
source
Location/Qualifiers
1..826
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5289949"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
```

BASE COUNT 147 a 252 c 240 g 187 t

## ORIGIN

Alignment Scores:  
 Pred. No.: 6,44e-95 Length: 826  
 Score: 1192.00 Matches: 249  
 Percent Similarity: 95.45% Conservative: 3  
 Best local Similarity: 94.32% Mismatches: 9  
 Query Match: 20.34% Indels: 7  
 DB: 12 Gaps: 1

US-09-397-967A-16 (1-1099) x BI557690 (1-826)

```

QY 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20
Db 37 ATGGCACTCCAGTGGAGAGACACCTCTGATCCCTCAGCGCTTTCAGCCTCTCATCC 96

QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40
Db 97 TCAGAGCAGGAGCCCTGTCATGTCTCTCTCCCGGGGACCTGGGCTCCCCAGCGA 156

QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaValAla 60
Db 157 TTGTCAATTC- TTGGGGAGTACTTGGCTGAGAGATTATGTGGCAGCTGCCAGGCC 215

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 216 TGTGGCATCTGCTGTTTATCATTCGCTTTCGCTCTGGCCACTGAGGACTTCTCTTGC 275

QY 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100
Db 276 TGGTTTCCCCCAGCCACATCTTCTGCATAGAGGACGTGGACACTCAAGTCTTGGTCTAC 335

QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120
Db 336 AGGCTACGCTTTTATTTCCCTGACTGTTGGGCTGGAGACATGTACCGCTT-GGGCTG 394

QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140
Db 395 CGCAAGATTG-ACCAGTGGCATCTTGTACTTACATGTTTGAACAATCTCTTTGCTCAG 453

QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160
Db 454 CACCGCAGTGACCTGGTGAGTGGGCGCTCCCGTGGGCTTAGCATGAAGGAGCAGGGA 513

QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180
Db 514 GAGTTCGTGAGCTGGCGCTGCTGGACTTGGCCAGATGGCTCTGTGACGAGCCCGAGCGC 573

QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200
Db 574 CCAGAGAGCTGCTGAGACGCTCAGTTACAAGCCTGTCTCGCGCCCGCAGCTCGCGAT 633

QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArg-IleArgArgThrValValLeuAl 220
Db 634 GTGATCCAGGCGCCAGAACTTCGTGACAGCGAGGACATCCGAGGACCGTGTCTTGGC 693

QY 220 aLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLe 240
Db 694 GCTGCG-CCGTGTGCTCGCTTGGAGGCGGACCGTACGCGCTCATGGCCAAAGTATTTCT 752

QY 240 uAspLeuGluArgLeuHisPro-AlaAlaThrThrGluThrPheArgValGlyLeuProG 260
Db 753 GGACTTGAAACGGTACATCCAGAGCGGCAACAACGAGA--CTCGTGTGGGGCTCCCGG 809

QY 250 lValaGln 262
Db 810 GCGCCAGG 817

RESULT 8
AK087255 2222 bp mRNA linear HTC 05-DEC-2002
LOCUS Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
DEFINITION library, clone:E030040D20 product:Janus kinase 2, full insert
sequence.
ACCESSION AK087255
VERSION AK087255.1 GI:26104150
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE Carninci, P. and Hayashizaki, Y.
AUTHORS High-efficiency full-length cDNA cloning
TITLE
```

```

JOURNAL MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kiteunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,
Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Sato,K., Schonbach,C., Seya,F., Shibata,Y., Storch,K.F., Suzuki,H.,
Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
and Hayashizaki,Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL MEDLINE 21085660
PUBMED 11217851
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL MEDLINE 21085660
PUBMED 11217851
REFERENCE 6
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE Direct Submission
JOURNAL MEDLINE 21085660
PUBMED 11217851
REFERENCE 7
AUTHORS Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,

```

```

URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
Location/Qualifiers
source
1..2222
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:E030040D20"
/db_xref="taxon:10090"
/clone="E030040D20"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
misc_feature
307..>2222
/note="Janus kinase 2 (MGD|G1:96629, GB|NM_008413,
evidence: BLASTN, 99%, match=1942)
putative"
BASE COUNT 633 a 493 c 547 g 549 t
ORIGIN
Alignment Scores:
Pred. No.: 2,26e-92 Length: 2222
Score: 1170.50 Matches: 261
Percent Similarity: 59.81% Conservative: 114
Best Local Similarity: 41.63% Mismatches: 225
Query Match: 19.97% Indels: 27
DB: 11 Gaps: 9
US-09-397-967A-16 (1-1099) x AK087255 (1-2222)
QY 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla 23
Db 358 CCTGTACATCAAGTAATGGTGTATATCTCTGGAAGTGTAAATCTGTGAAGCAGATAGAGCCA 417
QY 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43
Db 418 GTTCCTCAAGTGTATCTGTACCATCTCTCTGGCAAGCTGAAGGAGATCTGAAGTTT 477
QY 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla-CysGlyI 63
Db 478 CCAAGTGGAGAGTATGTTTCAGAGAAATTTTGTGTGCTGCTTCTTAAGCTTGATGAT 537
QY 63 eLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTyrPhePr 83
Db 538 TAGCGCTGTATCATATAATATGTTTCGTTAATAGTGAACCCGAAAGGATCTGTACCG 597
QY 83 oProSerHisLeuPheCysIleGluLeuValAspThrGlnValLeuValTyrArgLeuAr 103
Db 598 ACCCATCATGTCCTCCACATAGACAGTCACACAGGATGACATCTCTCAGGATAG 657
QY 103 gPheTyrPheProAspTyrPhePhe-----GlyLeuGluThrCysHisArgPheGlyLeuAr 121
Db 658 GTTCTACTTCCCTCATTTGTTGTTAGTGGCAGCAGCAGAACCTACAGATACGGAGTGC 717
QY 121 gLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHi 141
Db 718 CGGTGGGGCTGAAGCTCTCTGCTGACATGACATTTGTCTGATGCTTCTCTTTTGTCTCAGTG 777
QY 141 sArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGl 161
Db 778 GCGGATGATTTTGTTCACGGATGATAAAGTACTGTGATCATGATCAGGAGAA 837
QY 161 uPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPr 181
Db 838 GTGTCTTGGGTGGCGGTGTTAGCATGATGATGATAGTAAAGGAGAAAGACCACTCC 897

```

```

Qy 181 oGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProSerLeuArgAspVa 201
Db 898 ACTGGTGTCTATAACTCTGTACGTACAAGACATCTCTTACCAAGTCGGTTCGAGCGAA 957
Qy 201 lIleGlnGlyGlnAsnPheValThrArgArgArgileArgArgThrValValLeuAlaLe 221
Db 958 GATCCAGACTATCACATTTTAAACCCGGAAGCGAATCAGGTACAGATTTTCGCAGATTCAT 1017
Qy 221 uLeuProCysGlyLeuProGlyArgProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAs 241
Db 1018 TCAGCAATTCAGTCAATGTAAGCCACTGCCAGGAACCTAAACCTTAAGTATCTTATAAA 1077
Qy 241 pLeuGluArg-LeuHisProAlaAlaThrThrGluThrPheArgVal----- 256
Db 1078 CTTGGAACCCCTGCTGCTGCTCTTCTACACAGAACAGTTTGAAGTAAAGAAATCTGCAG 1137
Qy 257 --GlyLeuProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyI 276
Db 1138 GAGGT---CCTTCAGGTGAGGAGATTTTTCACCAACATTAATAAATCTGGAACCGTGGAA 1194
Qy 276 leProTrpSerSer-----AsnAspGluLeuPheGlnT 287
Db 1195 TTCAGTGTCTAAGAGGGAACATTAAGGAAGTGAACACTGACACAGACAGACGATACAGT 1254
Qy 287 hrPheCysAspPheProGluIleValAlaAspValSerIleAsnGlnAlaProArgValGlyP 307
Db 1255 TATATTGTGATTCCCTGATATATTGATGTCAGTATTAAAGCAAGCAACAG---GAAT 1311
Qy 307 roAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaG 327
Db 1312 GCTCAATGAAGTAGAATTGTAATCTGTCCTCAATAACCAAGATGTTAAAGTTTTCGAGATAG 1371
Qy 327 luPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgL 347
Db 1372 AACTTAGCTCATTAAGAACCCCTGTCATTCGTGTCATTAAATGACGGGTATTACAGAC 1431
Qy 347 euIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluG 367
Db 1432 TAACCTGGGATGGCACCATTACCTCTGCAAGAGGTGGCTCCCGCAGCTGTGCTCGAGA 1491
Qy 367 luGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysA 387
Db 1492 ACATACACAGCAACTGCCACGGCCCAATATCAATGGAATTTTGCCTAGCAAACTAAAGA 1551
Qy 387 laAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerP 407
Db 1552 AGCGGGGTAAACGACTGGACTATATGTCTACGATGCGCCCTTAAGGACTTCAACAAAT 1611
Qy 407 heLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleA 427
Db 1612 ACTTTCGACCTTGTCTGCTGAGCGAGAAATGTCATTGAATATAAACAACCTGTTTGATTA 1671
Qy 427 rgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuA 447
Db 1672 CGAAGAATGAGAAATGGAGAAATCAACCTCAGCGGCACTTAAGGAACTTCAGTAACCTTA 1731
Qy 447 rgGluLeuLeuAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrL 467
Db 1732 AGGACCTTTGAAATTTGCTTACCAGATGGAACCTGCGCTCAGACAGATATCATCTTCCAGT 1791
Qy 467 euThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgArg 487
Db 1792 TTACCAATGTGTCCTCCCAAGCAAGCAAGTAATCAACCTCTCTGCTCTCAGA---- 1847
Qy 487 lyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuT 504
Db 1848 --ACAAATGGTATTCTGATGTTTCAGATCTCACCAACATCTACAGAGCAATATAATGTA 1905
Qy 504 hrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyH 524
Db 1906 ATCAAAATGGTGTTCACAAAATCAGGAATGAAGATTTTAATATTAAAGATCTTGGCC 1965

```

524 isGlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp---GlyGluT 543  
1966 AAGGTACTTTTACAAAAATTTTAAAGTGTAAAGAGAGAAGTTGGAGATTATGGTCAAC 2025

543 hrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluS 563  
2026 TGCACAAAACGGAAGTCTTTTGAAGTCTCTAGATAAAGCAGACATAGAACCTATTTCAGAGT 2085

563 erPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuH 583  
2086 CTTTCTTCGAGCAGCAGCAGCATGATGAGTCAGCTTCTCACAACATTTGGTTTGAATT 2145

583 isGlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyA 602  
2146 ATGGTGTCTGTCTGTGGAGAGGAGAACATCTCTGGTTCAAGATTTGTAATAATTGGAT 2205

602 lalleAspMetTyrLeu 607  
2206 CACTGGATACATACCTG 2222

RESULT 9  
BM935182 698 bp mRNA linear EST 13-MAR-2002  
LOCUS UI-M-BH3-azq-f-08-0-UI.r1 NIH\_BMAP\_M\_S4 Mus musculus cDNA clone  
DEFINITION UI-M-BH3-azq-f-08-0-UI 5', mRNA sequence.  
ACCESSION BM935182  
VERSION BM935182.1 GI:19394334  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 698)  
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548

CONTACT: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: m5stemail.nih.gov  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 REVERSE.

Location/Qualifiers  
1..698  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH\_BMAP\_M\_S4"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not 1; Site 2: Eco RI; The  
NIH\_BMAP\_M\_S4 library is a subtracted library of a series,  
ultimately derived from a mixture of individually tagged  
(normalized) libraries from ten regions of the mouse brain  
(cerebellum, brain stems, olfactory bulbs, hypothalamus,  
cortex, amygdala, basal ganglia, pineal gland, striatum,  
hippocampus) after a series of subtractions to reduce the  
representation of cDNAs from which ESTs had already been  
generated. The following serially subtracted libraries  
were generated in this process: NIH\_BMAP\_M\_S4,

NIH BMAP M S3.3, NIH BMAP M S3.2, NIH BMAP M S3.1, NIH BMAP M S2, NIH BMAP M S1. The subtracted library (NIH BMAP M S4) was constructed as follows: PCR amplified cDNA inserts from NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH BMAP M S3.3, NIH BMAP M S3.2, and NIH BMAP M S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH BMAP M S4 library. This procedure has been previously described (Bonaldi, Lennon and Soares, Genome Research 6:791-806, 1996).

BASE COUNT 132 a 211 c 215 g 138 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1,62e-90 Length: 698  
Score: 1141.00 Matches: 225  
Percent Similarity: 96.15% Conservative: 0  
Best Local Similarity: 96.15% Mismatches: 6  
Query Match: 19,47% Indels: 4  
DB: 12 Gaps: 2

US-09-397-967A-16 (1-1099) x BM935182 (1-698)

QY 792 SerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAla 811  
Db 3 AGTCCTCGAGATGAGCTGTGC---GGTGGCGCCAGCTCTATGCTGCCAGGACCCCGC 59  
QY 812 IlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySer 831  
Db 60 ATATTCGAGAGAGACACCTTAACTACATCTCTTTGCTGGCAAGGCACTTTGGCAGC 119  
QY 832 ValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLys 851  
Db 120 GTGGAGCTGCGCGCTATCACCCCTGGGGACAATACGGGACCCCTGGTGGCAGTGAA 179  
QY 852 GlnLeuGlnHisSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnLeu 871  
Db 180 CAGCTACACACAGCGGGCCAGACACAGCAGAGGAGACTTCAGCGGAGATTACAGATCCT 239  
QY 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891  
Db 240 AAGGCTCTGCACAGCAGCTTATCTGTCAGTACCGGGAGTAGCTATGGCCAGGTGCG 299  
QY 892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGln 911  
Db 300 CAGAGCTCGGTTGGTGATGGAGTACCTGCCAGCGGCTGCTGGAGACTTCCTGCAG 359  
QY 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnLeuCysLys 930  
Db 360 CCCCATCGCGCGCGCTGCACACCGACCCCTACTGCTGCTTGGCAGATCTGCAAG 419  
QY 931 GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIle 950  
Db 420 GGCATGGAGTACTTGGTGGCGCGCTGCTGACACCGTGAGCTGCTGCGCGCAACATC 479  
QY 951 LeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuPro 970  
Db 480 TTGGTGGAGCAGAGGCTCATGTGAAGATCGGGACTTCGGCTCGTAAAGTGTGCTGCC 539  
QY 971 LeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIle-PheTyrTyrAl 990  
Db 540 CTGGGAAGAGACTACTACGTGGTCCGCGAGCTGCGCAAGCCCATCTTNTTGGTATGC 599  
QY 990 aProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyVa 1010  
Db 600 CCCGAGTCCCTATCTGCAC-ATCTTCTCNCGCCAATCTGACGTGTGGAGCTTCGGAGT 658  
QY 1010 lValLeuTyrGluLeuPheThrTyrCysAspLysSerCys 1023

Db 659 GGTTGTTGTACGAGCTCTTACCTACTGCGACAGAGCTGC 698

## RESULT 10

BF150250

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BF150250 675 bp mRNA linear EST 29-DEC-2000  
uy83e09.y1 NCI CGAP Mam5 Mus musculus cDNA clone IMAGE:3666184 5,  
similar to TR:P97423 P97423 JANUS KINASE 3 ; mRNA sequence.

BF150250 GI:11031645

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 675)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished

Other ESTs: uy83e09.x1

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[image.llnl.gov/image/html/iresources.shtml](http://image.llnl.gov/image/html/iresources.shtml)

MGI:1426952

Seq primer: -40RP from Gibco

High quality sequence stop: 390.

Location/Qualifiers

source

1. .675

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="IMAGE:3666184"

/tissue\_type="tumor, gross tissue"

/dev\_stage="7 months"

/lab\_host="DH10B"

/clone\_lib="NCI-CGAP\_Mams"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Library constructed by Life Technologies. Investigators

providing samples: Lothar Hennighausen/Robin Humphreys,

NIH"

BASE COUNT 138 a 215 c 197 g 123 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 6,41e-90 Length: 675

Score: 1134.00 Matches: 217

Percent Similarity: 96.02% Conservative: 0

Best Local Similarity: 96.02% Mismatches: 7

Query Match: 19,35% Indels: 2

DB: 10 Gaps: 2

US-09-397-967A-16 (1-1099) x BF150250 (1-675)

QY 705 GlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeu 724

Db 1 GGCTTTGGAGCCACCACCTGGGAGGTGTTTCAGCGGGGACCCGCCACATCACCTCGCTG 60

QY 725 GluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLeuLysTrp 744

Db 61 GAGCCCGCCAAAAGCTGAAGTCTATGAGGACCGAGGACAGCTGCCCGCTCTCAATGG 120

QY 745 ThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgArgProSer 764

Db 121 ACAGAACTGGGGGACTTATCACACAGTCATGGCGTATGATCTGCGCGCGCCCTCC 180  
 Qy 765 PheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLeuSer 784  
 Db 181 TTCCGAGCTATCTCTCAGAGACCTCAACGGCTCATTTACATCAGATTACGAGCTCTCTCA 240  
 Qy 785 AspProThrProGlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeu 804  
 Db 241 GACCCACACTGGCATCCCGAGTCCTCGAGATGAGCTGTGC-----GGTGGCGCCACGCTC 297  
 Qy 805 TyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuIleTyrIleSerLeuLeu 824  
 Db 298 TATGCTTCCAGGACCCCGCCATATTTCGAGGAGACACCTTAAAGTACATCTCTTTGCTG 357  
 Qy 825 GlyIleGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeuGlyIleAspAsnThr 844  
 Db 358 GCGAAGGCGCACTTTGGCAGCGTGGAGCTGTGCGCGCTATGACCCCTGGGGGACAAATACG 417  
 Qy 845 GlyProLeuValAlaValIleGlnLeuGlnHisSerValProAspGlnGlnArgAspPhe 864  
 Db 418 GGACCCCTGTGGCAGTGGACACTACAGCAGCGGGCCGACACCGAGCAGAGGACTTC 477  
 Qy 865 GlnArgGluIleGlnIleLeuIleAlaLeuHisSerAspPheIleValIleTyrArgGly 884  
 Db 478 CAGCGGAGATTCAGATCCTTAAGCTCTGACAGCGACTTCATCGTCAAGTACCGGCGGA 537  
 Qy 885 ValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGly 904  
 Db 538 GTACGCTATGGCCGANGTCGCGAGCGCTGGCGTGGTGTGATGAGTACCTCCCGAGCGGC 597  
 Qy 905 CysLeuArgAspLeuGlnArgHisArgGly--LeuHisThrAspArgLeuLeuLeu 923  
 Db 598 TGCCTGGAGACTTCTGACGGCCATCGCGCGCTGACACCGACCGCGCTACTGCTG 657  
 Qy 924 PheAlaTrpGlnIleCys 929  
 Db 658 TTCGCTGGCAGATCTGC 675

RESULT 11  
 BY730320  
 LOCUS  
 DEFINITION BY730320 RIKEN full-length enriched, 13 days embryo heart Mus  
 Musculus cDNA clone D330022C11 5', mRNA sequence.  
 BY730320  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus (house mouse)  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 677)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,  
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,  
 Gajober, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,  
 Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,  
 Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani,  
 L.E., Cousins, S., Dalia, E., Dragani, I.A., Fletcher, C.F., Forrest,  
 A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,  
 Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,  
 King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,  
 P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,  
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,  
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.D., Qi, D.,  
 Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring,  
 B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou,  
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale,  
 R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,  
 Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa,  
 M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,  
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,  
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.

FEATURES  
source

Location/Qualifiers  
 1..677  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="D330022C11"  
 /tissue\_type="heart"  
 /dev\_stages="13 days embryo"  
 /lab\_host="DH10B"  
 /clone\_lib="RIKEN full-length enriched, 13 days embryo heart"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5].  
 GAGAGAGAGCGCGCGCACTCGACTTTTCTTTTCTTTT 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5].  
 GAGAGAGATTCGAGTTAATTAATTAATTCCTCCCTCCCTCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLIC 1."  
 117 a 203 c 188 g 169 t

BASE COUNT  
 ORIGIN

, Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 JOURNAL  
 MEDLINE 22354683  
 PUBMED 12466851  
 COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp/  
 URL: http://genome.gsc.riken.go.jp/  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct  
 Submission  
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

## Alignment Scores:

Pred. No.: 3,13e-85 Length: 677  
Score: 1081.00 Matches: 209  
Percent Similarity: 99.52% Conservative: 0  
Best Local Similarity: 99.52% Mismatches: 1  
Query Match: 18.45% Indels: 1  
DB: 14 Gaps: 0

US-09-397-967A-16 (1-1099) x BY730320 (1-677)

QY 1 MetaLProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSer 20  
Db 47 ATGGCACTCCAAAGTGAGGAGACACCTCTGATCCCTCAGCGCTTGTGAGCCTTCATCC 106  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
Db 107 TCAGAGGCGAGGAGCCCTGATGCTTCCTCCCGGGGAGCCTGGGCTCCCGAGCA 166  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
Db 167 TTGTCTATCTCTTTGGGAGCTACTTGGCTGAGATTATGTGTGAGCTGCCAAGGCC 226  
QY 61 CysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
Db 227 TGTGGCATCTCGCTCTTTATCTATCTGCTTTTCGCTCGCCACTGAGACTTCTCTTG 286  
QY 81 TrpPheProProSerHisLePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
Db 287 TGGTTTCCCCCAAGCCACATCTTCTCATAGAGGAGCTGACACTCAAGTCTTGCTAC 346  
QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
Db 347 AGCTACGCTTTATTTCTCTGACTGTTGGCTGGAGACATGTCAACGCTTTGGGCTG 406  
QY 121 ArgLysAspLeuThrSerAlaLeuLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
Db 407 CGCAAGATTTCACCAAGTCCCTGCTGACTTACAGCTTTAGAACATCTCTTTGCTCAG 466  
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
Db 467 CACCGCAGTGACTGCTGAGTGGCGCCCTCCCGTGGGCTTAGCATGAAGGAGCGGA 526  
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
Db 527 GAGTTCCTGAGCTGGCCGTGTGACTTGGCCACAGTCTGCTGAGCAGGCCCGCC 586  
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
Db 587 CCAGGAGAGCTGCTGAAGACGGTCACTTACAAGCCTGTCTGCCGCCAGCCTGCGCAT 646  
QY 201 ValIleGlnGlyGlnAsnPheValThrArg 210  
Db 647 GTGAT-CAGGGCCAGAACTTCGTGACACGC 675

## RESULT 12

BX370312

LOCUS BX370312 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
DEFINITION Homo sapiens cDNA clone CS0DJ012Y019 5-PRIME, mRNA sequence.

## ACCESSION

BX370312

VERSION BX370312.1 GI:30453895

## KEYWORDS

EST.

## SOURCE

Homo sapiens

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 927)

Li, W. B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0BAG043ZB12\_CS04122\_1.

## FEATURES

## source

1..927  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DJ012Y019"  
/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/cell\_line="JURKAT"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT  
10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."  
BASE COUNT 159 a 298 c 295 g 167 t 8 others  
ORIGIN

## Alignment Scores:

Pred. No.: 3.47e-84 Length: 927  
Score: 1071.50 Matches: 230  
Percent Similarity: 82.18% Conservative: 19  
Best Local Similarity: 75.93% Mismatches: 49  
Query Match: 18.28% Indels: 9  
DB: 13 Gaps: 2

US-09-397-967A-16 (1-1099) x BX370312 (1-927)

QY 82 PheProProSerHisLePheCysIleGluAspValAspThrGlnValLeuValTyrArg 101  
Db 5 TTTCCCCCGGACCAATNCTTCTCGTGGGGNAGCCAGCAGCCCAAGTCTCTGTACAG 64  
QY 102 LeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeuArg 121  
Db 65 ATTCCGCTTTACTTCCCCAT-TGGTTTGGGCTGGAGAGTGGCCACCGCTTCGGGCTACGC 123  
QY 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141  
Db 124 AAGGATTTGGCCAGTGTCTATCTTGACCTGCCAGTCTCGAGCACCTCTTGGCCAGCAC 183  
QY 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 161  
Db 184 CGCAGTGAAGTCTGAGTGGGCGCTCCCGTGGGCTCAGTCTCAAGGAGCAGGGTGAG 243  
QY 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181  
Db 244 TGTCTCAGCTGGCGGTGTGGACCTGGCCCGAGTGGCGGAGAGCAGGCCCGCGCGCG 303  
QY 182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201  
Db 304 GGAGAGCTGCTGAAGACTGTCTAGCTACAGGCTGTCTACCCCAAGCTTCGGGAGCTG 363  
QY 202 IleGlnGlyGlnAsnPheValThrArgArgIleArgArgThrValValLeuAlaLeu 221  
Db 364 ATCCAGGGCTGAGCTTCTGAGCGGGAGCGCTATTTCGAGGAGCGGTGCGCAGAGCCCTG 423  
QY 222 LeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeuAs 241  
Db 424 CG-CCCGTGGCGGCTGCCAGGACACCGGCACCTCGCTCATGGCAAGTACATCATGA 482  
QY 241 pLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAl 261  
Db 483 CTGGAGCGCTGGATCCAGCGGGCGGCCCGAGACCTTCCACGTGGGCTCTCTTGGGGC 542  
QY 261 a-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTr 278  
Db 543 CTTGTGTGCCACGACGGGCTGCTCGCGTGGTGTGTGACGGCGGATCGCCTG 602  
QY 278 pSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAspVa 297



```

Db 603 GACCCAGGGAGACAGAGGTCTCCAGCCCTCTCGAGATTCAGAGAACTCGTAGCAT 662
Qy 297 lserlIasnglnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValTh 317
Db 663 TAGCATCAAGCAGGCCCGCGCTTGGCGCGCGGAGAGACCGCTGCTGCTACTGTTAC 722
Qy 317 rAtgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPh 337
Db 723 CAGGACACACACAGATTTTAGAGGCGGAGTTCCTCAAGGCTGCCGAGGCTCTGTGNT 782
Qy 337 eValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysLy 357
Db 783 CGTGGCGCTCGTGACGCGTACTTCGGCTGACACGAGCTCCAGCAGCTTCTTCTCGAG 842
Qy 357 sGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleTh 377
Db 843 GAGGTTTGCACCGCGCGAAGCTGTG-NAGGAAAGTGCCNANCAGTGCCA-GGNCCCCAC 900
Qy 377 rLeuAsp 379
Db 901 TCTGGAC 907

```

## RESULT 13

BY746452

LOCUS

DEFINITION

BY746452

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BY746452 655 bp mRNA linear EST 17-DEC-2002  
 cells (NOD) Mus musculus cDNA clone E430014A10 5', mRNA sequence.  
 BY746452  
 BY746452.1 GI:27174142  
 EST.  
 Mus musculus (house mouse)  
 Mus musculus  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 655)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nikaado, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,  
 Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,  
 Gojobori, T., Baidarelli, R., Hill, D. P., Bulc, C., Hume, D. A.,  
 Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S.,  
 Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V., Chocthia, C., Corbani,  
 L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest,  
 A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,  
 Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J.,  
 Jarvis, E. D., Kanaï, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M.,  
 King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,  
 P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki,  
 H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,  
 Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D.,  
 Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,  
 M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,  
 R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,  
 Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa,  
 M., Yang, I., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,  
 Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura,  
 M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,  
 Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, T., Shibata,  
 K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,  
 E. S., Rogers, J., Birney, E., and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 22354683  
 12466851  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-2 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222

Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda,  
 S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K.,  
 Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno,  
 H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K.,  
 Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y.,  
 Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct  
 Submission

Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
 Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
 Trust/MRC Building Addenbrookes Hospital Cambridge) whose  
 assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

## FEATURES

source

Location/Qualifiers  
 1..655  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NOD"  
 /db\_xref="taxon:10090"  
 /clone="E430014A10"  
 /tissue\_type="thymus"  
 /cell\_type="thymic cells"  
 /clone\_lib="RIKEN full-length enriched, 2 days neonate  
 thymus thymic cells (NOD)"

BASE COUNT 111 a 197 c 180 g 167 t

## ORIGIN

## Alignment Scores:

Pred. No.: 6,42e-81 Length: 655  
 Score: 1032.00 Matches: 195  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 17.61% Indels: 0  
 DB: 14 Gaps: 0

US-09-397-967a-16 (1-1099) x BY746452 (1-655)

Qy 1 MetalProProSerGluGluThrProleulleProGlnArgSerCysSerLeuSerSer 20  
 Db 63 ATGCACCTTCAAGTGAGGACACCTCTGTATCTCTCAGCGCTCTTGACGCTCTTCATCC 122  
 Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProProGlnArg 40  
 Db 123 TCAGAGGAGGAGCGCTGATGTCTCTCTCTCCCGGGACCTGGGCTCCCGAGCGA 182  
 Qy 41 LeuSerPheSerPheGlyAspTyrleuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
 Db 193 TTGTCACTCTCTTTGGGGACTACTTGCTGAGGATTTATGTGCGAGCTGCCAAGGCC 242  
 Qy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
 Db 243 TGTGGCATCTCTGCTGTTTATCTTCGCTTTTGGCTTGGCCACTGAGACTTCTCTTGC 302



QY 81 TtpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
 Db 303 TGGTTTCCCCCAAGCCACATCTTCTGCATAGAGGAGCTGGACACTCAAGTCTTGGTCTAC 362  
 QY 101 ArgLeuArgPheTyrPheProAspTrrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
 Db 363 AGCTACGCTTTATTTCCCTGACTGGTTTGGGCTGGAGACATGTCCACCGCTTTGGGCTG 422  
 QY 121 ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
 Db 423 CGCAAGATTGACCATGGCCATCTTGTGCTTACGCTTTTGAACATCTCTTTGTCTCAG 482  
 QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
 Db 483 CACCGCAGTGACCTGTGTAGTGGCGCTCCCGGTGGGCTTAGCATGAAGAGCAGGGA 542  
 QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
 Db 543 GAGTTCTGAGCCGTGGCGCTGTGGACTTGGCCCATGTGGCCAGATGGCTGAGCAGGCCAGCGC 602  
 QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuPro 195  
 Db 603 CCAGGAGAGCTGCTGAAGACGGTCAGTTACAAAGCCTGTCTGCCG 647

RESULT 14  
 BG870513  
 LOCUS 602791432F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4922742 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG870513  
 VERSION BG870513.1 GI:14221053  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 907)  
 NIH-MGC <http://imgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
 Tissue Procurement: Jeffrey E. Green, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLAM10842 row: p column: 07  
 High quality sequence stop: 742.  
 Location/Qualifiers  
 1..907  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:4922742"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NCI\_CGAP\_SG2"  
 /note="Organ: salivary gland; Vector: pCMV-SPORT6; Site 1:  
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo  
 dT. Average insert size 1.3 kb. Constructed by Life  
 Technologies. Note: this is a NCI\_CGAP Library."  
 166 a 279 c 275 g 185 t 2 others

## FEATURES

source

Alignment Scores:  
 Pred. No.: 1,16e-80 Length: 907  
 Score: 1031.50 Matches: 216  
 Percent Similarity: 88.93% Conservative: 1  
 Best Local Similarity: 88.52% Mismatches: 23  
 Query Match: 17.60% Indels: 7

## BASE COUNT

ORIGIN

DB: 12 Gaps: 3  
 US-09-397-967A-16 (1-1099) x BG870513 (1-907)  
 QY 858 ProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAsp 877  
 Db 39 CCAGACCAGCAGAGGAGACTCCAGCGGAGATTCCAGATCTCTTAAGGCTCTGCACAGCGAC 98  
 QY 878 PheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuVal 897  
 Db 99 TTCATCGTCAAGTACACGGGAGTCAGCTATGGCCAGGTCGCCAGAGCCCTCGCGTTGGTG 158  
 QY 898 MetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly---Leu 916  
 Db 159 ATGGAGTACTGCCAGCGGCTGCCAGACTTCTGCGAGCCCATCGCGCGCGCTG 218  
 QY 917 HisThrAspArgLeuLeuPheAlaTrrPheGlnIleCysLysGlyMetGluTyrLeuGly 936  
 Db 219 CACACCGACCGCCTACTGCTGTTCGCTGGCAGATCTGCAAGGGCATGGAGTACCTGGGT 278  
 QY 937 AlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAla 956  
 Db 279 GCGCGCGCTGCGTACACCGAGACCTGGTGGCGCGCAACATCTTGTGGAGAGCGAGGCT 338  
 QY 957 HisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyr 976  
 Db 339 CATGTGAAGATCGCGACTTCGGCTCGCTAAGCTGCTGCCCTGGGAAGGACTACTAC 398  
 QY 977 ValValArgGluProGlyGlnSerProIlePheTrrPyrAlaProGluSerLeuSerAsp 996  
 Db 399 GTGGTCCGCGAGCCTGGCCAAAGCCCATCTTTTGGTATGCCCGCGAGNTCTTATCTGAC 458  
 QY 997 AsnIlePheSerArgGlnSerAspValTrrPheGlyValValLeuTyrGluLeuPhe 1016  
 Db 459 AACATCTTCTCCGCGCAATCTGACGTGGAGCTTCGGAGTGGTGTGTACGAGCTCTTC 518  
 QY 1017 ThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGlu 1036  
 Db 519 AC-TACTGGCAAGAGCTGCGAGCCATCCGCTGAGTTCCTGAGCATGATGGGCGCTGAG 577  
 QY 1037 ArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgLeuPro 1056  
 Db 578 CGTGAAGGACCCCGTTCTTGGCTC-CTGAGGCTCTGGCAGAGGCCGAGC-CTCCCA 635  
 QY 1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076  
 Db 636 CATCTTTCACCTGCC---CACGAGGTTTCAGAGCTCATGAGCTGTGTCTGGCGCCGAGC 692  
 QY 1077 ProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTrrArgGly 1096  
 Db 693 CCG---ACGACGGGAGCGCTTCGGACCTTCAGCCCTGAGCCCCCAGCTGGACCGCTGTGGAGTGA 749  
 QY 1097 -ArgProGly 1099  
 Db 750 AAGACCCGGG 759  
 RESULT 15  
 BG873355  
 LOCUS 602794328F1 NCI\_CGAP\_SG2 Mus musculus cDNA clone IMAGE:4925598 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG873355  
 VERSION BG873355.1 GI:14223895  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 720)  
 NIH-MGC <http://imgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.

Email: cgapbs-x@mail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LHAM10850 row: 9 column: 07  
High quality sequence stop: 709.  
Location/Qualifiers  
1. .720  
source  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4925598"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI CGAP SG2"  
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1:  
Not1; Site 2: Sal1; Cloned unidirectionally. Primer: Oligo  
dT. Average insert size 1.3 Kb. Constructed by Life  
Technologies. Note: this is a NCI CGAP Library."  
BASE COUNT 123 a 239 c 229 g 129 t  
ORIGIN

## Alignment Scores:

Pred. No.:	1.51e-80	Length:	720
Score:	1028.50	Matches:	201
Percent Similarity:	95.71%	Conservative:	0
Best Local Similarity:	95.71%	Mismatches:	8
Query Match:	17.55%	Indels:	3
DB:	12	Gaps:	1

US-09-397-967A-16 (1-1099) x BG873355 (1-720)

Qy	891	ArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeu	910
Db	1	CGCCAGAGCCCTGGGTTGGTGTGATGGAGTACTTGCACCGCGCTGCTCGGAGACTTCCTG	60
Qy	911	GlnArgHisArgGly--LeuHisThrAspArgLeuLeuPheAlaTrpGlnLeuCys	929
Db	61	CAGCGCCATCGCGCGCGCTGCACACCGCGCTACTGCTGCTTGGCTGGCAGATCTGC	120
Qy	930	LysGlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsn	949
Db	121	AAGGGCATGGAGTACTCGGGTGGCGCGCTGCGTACACCGTACCTGGCTGGCGGCAAC	180
Qy	950	IleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeu	969
Db	181	ATCTTGGTGGAGAGCGAGGCTCATGTGAAGATCGGGACTTCGGCTCGCTAAGCTGCTG	240
Qy	970	ProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyr	989
Db	241	CCCTGTGGAAAGGACTACTACGTGTGCGGAGCTGGCCAAAGCCCATCTTTTGGTAT	300
Qy	990	AlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGly	1009
Db	301	GCCCGGGAGTCCCTATCTGACAAACATCTTCCCGCCCAATCTGACGTGGAGCTTCGGA	360
Qy	1010	ValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPhe	1029
Db	361	GTGGTGTGTACGAGCTCTTCACCTACTTCGACACAGAGCTGCAGCCCATCCGCTGAGTTC	420
Qy	1030	LeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuLeuLeu	1049
Db	421	CTGCGCATGATGGGCGCTGAGCGTGAAGGACCCCGCTCTGCGCGCTCTCGAGCTGCTG	480
Qy	1050	AlaGluGlyArgArgLeuProProProThrCysProThrGluValGlnGluLeuMet	1069
Db	481	GCAGAGGGCGGACGGCTCCACCAACTCCACCTCCCGCCACCGAGTTTCAGGAGCTCATG	540
Qy	1070	GlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGln	1089

Db	541	CAGCTGTGTGGGCGCCAGCCCGCA-GACCGGCGAGCTTCGG-ACCTGAGCCCCCAG	598
Qy	1090	LeuAspProLeuTrpArgGlyArgProGly	1099
Db	599	CTGGACGCGCTGTGGCTGGAAGACCCGGA	628

Search completed: February 4, 2004, 08:14:50  
Job time : 6493.54 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 01:30:36 ; Search time 172.373 Seconds  
(without alignments)  
2814.136 Million cell updates/sec

Title: US-09-397-967A-16

Perfect score: 5860

Sequence: 1 MAPSEETPLIPQSCSLSS.....RPAFATLSPQLDPLWRGREG 1099

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09397967/runat 03022004.175632.28596/app\_query.fasta.1.1486  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINWATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pot -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967 @CGN 1.1.81 @runat 03022004.175632.28596 -NCPU=6 -ICPU=3  
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5223	89.1	4016	5	PCT-US95-08354A-1
2	4597	78.4	3620	4	US-09-016-434-297
3	4418	75.4	3807	2	US-08-357-598-1
4	4418	75.4	3807	2	US-08-003-289-1
5	4418	75.4	3807	5	PCT-US95-16435-1
6	2629	44.9	3435	3	US-09-046-158A-21
7	2621	44.7	4482	2	US-08-567-508C-1
8	2621	44.7	4482	3	US-09-196-480-1
9	2621	44.7	5117	4	US-09-972-800A-15
10	2619.5	44.7	3629	1	US-08-097-997A-8
11	2619.5	44.7	3629	3	US-08-665-574C-8
12	2619.5	44.7	3629	3	US-08-946-994-8

13	2594	44.3	4078	4	US-09-016-434-297
14	2360	40.3	3495	1	US-08-446-038B-2
15	2360	40.3	3495	1	US-08-446-010B-2
16	2360	40.3	3495	1	US-08-805-445-2
17	2360	40.3	3495	2	US-08-064-067D-2
18	2360	40.3	3495	2	US-09-066-208-2
19	2360	40.3	3495	4	US-08-980-080-3
20	1904	32.5	3429	1	US-08-097-997A-10
21	1904	32.5	3429	3	US-08-665-574C-10
22	1904	32.5	3429	3	US-08-946-994-10
23	1904	32.5	4234	1	US-08-446-038B-1
24	1904	32.5	4234	1	US-08-446-010B-1
25	1904	32.5	4234	1	US-08-805-445-1
26	1904	32.5	4234	2	US-08-064-067D-1
27	1904	32.5	4234	2	US-09-066-208-1
28	1817.5	31.0	4080	4	US-09-016-434-1353
29	1817.5	31.0	4176	4	US-09-972-800A-17
30	1815.5	31.0	3561	1	US-08-097-997A-12
31	1815.5	31.0	3561	3	US-08-665-574C-12
32	1815.5	31.0	3561	3	US-08-946-994-12
33	482	8.2	3845	2	US-08-220-240A-4
34	475	8.1	4530	1	US-08-229-515A-9
35	475	8.1	4530	1	US-08-645-865-9
36	475	8.1	4530	4	US-09-167-322-4
37	475	8.1	4530	4	US-09-527-487-1
38	475	8.1	4530	4	US-09-877-177A-11
39	474	8.1	3768	2	US-08-625-101-1
40	474	8.1	3768	2	US-08-356-786-1
41	474	8.1	4473	2	US-09-048-804-1
42	474	8.1	4473	3	US-09-056-105-26
43	466.5	8.0	5532	2	US-08-475-035-3
44	466.5	8.0	5532	4	US-09-676-610B-17
45	465.5	7.9	5484	3	US-09-632-580A-3

## ALIGNMENTS

RESULT 1  
PCT-US95-08354A-1  
; Sequence 1, Application PC/TUS9508354A  
; GENERAL INFORMATION:  
; APPLICANT: Temple University - Of The  
; APPLICANT: Commonwealth System of Higher Education  
; TITLE OF INVENTION: JAK3 PROTEIN TYROSINE  
; TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seidel, Gonda, Lavorigna  
; ADDRESSEE: & Monaco, P.C.  
; STREET: Suite 1800, Two Penn Center  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/08354A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/272,368  
; FILING DATE: 8 July 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 6056-203 PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549



QY 421 TyrIysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSer 440  
Db 1921 TACAAGGGCTGCTCATCCGCCAGAGACCCAGCGGGGCTTTCTCCCTGGTTGC-CTCAGC 1979  
QY 441 GlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgVal 460  
Db 1980 CAGCCCCACAGAAAGCTCGCGGAGCTGCTTGGACG-TGCTGGAAATTC-GGGCTGCGAGTA 2037  
QY 461 AspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProIysGlyLysSerAsn 480  
Db 2038 GACGGTGTCTGCTGAACTTACATCTGCTGGCTCCAGACCCAGGAAAGTCCAAAT 2097  
QY 481 LeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCys 500  
Db 2098 TTGATCGTGTGCGAAGGGGCTGCACCCCGCGCTGCCCTGGCTGCTCCCGTCTGTC 2157  
QY 501 CysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGlu 520  
Db 2158 TGTGGCTGACACAGCTGAGCTTCCACAAATTCACAGCGACAGCGCTG--GGACACGAG 2214  
QY 521 AsnLeuGlyHisGlySerPheThrIysIlePheArgGlySerArgArgGluValValAsp 540  
Db 2215 AACCTGGGTTCACGGTTCCTTTACCAAGATCTTCCGTGGCGCAGCGGGAGGTGCGGAT 2274  
QY 541 GlyGluThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCys 560  
Db 2275 GGTGAGACACATGACTCGGAAGTCTCTCTGAAGTTCATGAGTCTCCAGACATCGGAATGC 2334  
QY 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580  
Db 2335 ATGGAGTCTTTTCTGGAGCCGCAAGCTTGATGAGCCAAATTCCTACCGCACCTGGTG 2394  
QY 581 LeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGlnGluPheValTyrLeu 600  
Db 2395 TTACTGCACGGCTGTGCATGGCTGGAGACACATCATGGTGCAGAAATTTGTGTATCTA 2454  
QY 601 GlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeu 620  
Db 2455 GGAGCAATTGACATGACTACCTGCGCAAGCGTGCCACCTGGTGTGACCCAGCTGGAACTG 2514  
QY 621 GlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis 640  
Db 2515 CAGTGACCAAGCAGCTGGCATATGCCCTTAATCTTGGAGGACAAAGGCGCTTCTCTCAC 2574  
QY 641 GlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyValAspGlyAsnProPro 660  
Db 2575 GGCAAGGCTTCAGCACGGAAGGTGCTCTGGCTCTGAGGGGGGTGATGGGAATCCACCT 2634  
QY 661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680  
Db 2635 TTCATTAAAGCTGAGTGATCTGCTGCTCAGTCCCTGCTGAGGCTGGAAATGCTCACC 2694  
QY 681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700  
Db 2695 GACAGAAATACCTGGGTGGCCCCGCAAGTCTTCCAGGAGGCTCAGACATCTGCTTGGAG 2754  
QY 701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720  
Db 2755 GCTCACAAGTGGGCTTTGGAGCCACCATGCTGGAGGTGTTACGGGGGGGACCCGCGCCAC 2814  
QY 721 IleThrSerLeuGluProAlaLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740  
Db 2815 ATCACCCTGGCTGGAGCCCGCCCAAAAGCTGAAGTCTTATGAGGACCCAGGACAGCTGCC 2874  
QY 741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760  
Db 2875 GCTCTCAATGGACAGAACTGGCGGAGCTTATCAGACAGTGCATGGGGGTATGATCTCTGGC 2934  
QY 761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780  
Db 2935 CGGGCGCCCTTCTCCAGCTATCTCTCAGAGACCTCAACGCCCTCATATCATCATGATTAC 2994  
QY 781 GluLeuLeuSerAspProThrProGlyIleProSerProArgAspGluLeuCysValAla 800

Db 2995 GAGCTCTCTCA-GACCCACACCTGGCATCCCGAGTCTCTGAGATGAGCTGTGCGT---- 3049  
QY 801 GlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLysTyr 820  
Db 3050 GGGCGCCAGCTCTATGCTTGCAGGACCCCGCATATC-GAGGAGAGACACCTTAAGTAC 3108  
QY 821 IleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspProLeu 840  
Db 3109 ATCTCTTTTCTGGCAAGGGCAACTTGGCAGCGTGGAGCTGTGCGCTATGACCCCTG 3168  
QY 841 GlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAspGln 860  
Db 3169 ---GACATATCCGGACCCCTGGTGGCAGTGAACAGCTACAGCACAGCGGGCCAGACCA 3225  
QY 861 GlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIleVal 880  
Db 3226 CAGAGGGACTTCCAGCGGAGATTTCAGATCTTAAAGCTCTGCACAGCGACTTCATCGTC 3285  
QY 881 LysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGluTyr 900  
Db 3286 AAGTACCGGGAGTCAGCTATGGGCAGCTGCCAGAGCGCTCGGTGGTGTGATGAGTAC 3345  
QY 901 LeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThrAsp 919  
Db 3346 CTGCCCAGCGGCTGCTGGAGACTTCTGACAGCCCATCTGCGCGGCCCTGACACACCGAC 3405  
QY 920 ArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArgArg 939  
Db 3406 CGCCTACTGCTTTCGCTTGGCAGATCTGCAAGGCGATGGAGTACTCTGGTGGCGCGCGC 3465  
QY 940 CysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisValLys 959  
Db 3466 TCGGTACACCGGTGACTGGCTGCGCGCAACATCTTGGTGGAGAGCGAGCTCATGTGAG 3525  
QY 960 IleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValValArg 979  
Db 3526 ATCGCGGACTTTCGGCTCGTAAAGCTGCTGCCCTGGGAAAGGACTACTACGTGTCGCG 3585  
QY 980 GluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsnIlePhe 999  
Db 3586 GAGCTGSCCAAAAGCCCATCTTTTGGTATGCCCGGAGTCCCTATCTGACAAATCTTC 3645  
QY 1000 SerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyrCys 1019  
Db 3646 TCCCGCCAAATCTGACGTGTGAGCTTCGGAGTGTGTGTACGAGCTCTTCACTACTGTC 3705  
QY 1020 AspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGluGly 1039  
Db 3706 GACAAAGAGCTCAGCCCATCCGCTGAGTTCCTGCGCATGATGGGCGCTGAGCGTGAAGA 3765  
QY 1040 ProProLeuCysArgLeuLeuLeuAlaGluGlyArgArgLeuProProProPro 1059  
Db 3766 CCCCCTGCTGCGGCTTCTGGAGCTGCTGGCAGAGGCGCCAGCCCTCCACACCTCTCC 3825  
QY 1060 ThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGluProHisAsp 1079  
Db 3826 ACCTGCCCCACCGAGGTTTCAGAGCTCATGCACTGTGCGTGGCGGCCAG-CGACACGAC 3884  
QY 1080 ArgProAlaPheAlaThrLeuSerProGluLeuAspProLeuTrpArgGlyArgPro 1098  
Db 3885 CGGCACGCTTGGCACCTTGGACCCCTGAGCCCCCAGCTGACGCGCTGTGGCGTGAAGACCC 3941

## RESULT 2

US-09-016-434-1055  
; Sequence 1055, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490

456	CGCAAGGATTGGCCAGTGCCTATCCTTGACCTGCCAGTCTCGAGCACTCTTTGGCCAG	515	HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly	160	161	GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg	180	516	CACCGCATGTGCTGGTGGGCGCTCCCGTGGCGCTCAGTCTCAAGGAGCAGGCT	575	166	GAGTGTCTCAGCCTGGCCGCTGTGGACCTGGCCGAGTGGCGGAGCAGGCGCCAGCGG	635	181	ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp	200	636	CGGGAGAGCTCTGAAGACTGTACAGCTACAGGCTGCTACCCCCAAGCCCTGCGGAC	695	201	ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla	220	696	CTGATCCAGGGCTGAGCTTCGTGACGCGAGGCGGTATTTCGAGGAGCGGTGCGCAGAGCC	755	221	LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLeu	240	756	CTGCG-CCGCGTGGCGCCTGCCAGCAGACCGGCACTCGCTCATGCCAAGTACATCAT	814	240	uAspLeuGlnArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGly	260	815	GGACCTGGAGCGGCTGGATCCAGCCGGGCGCCGAGAGCTTCCACGTGGGCTCCCTGG	874	260	yAla-----GlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr	277	875	GGCCCTTGGTGGCCACAGCGGGCTGGGGTGTCTCCGCGTGGCTGGTACGCGCGCATCGC	934	277	oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs	296	935	CTGGACCCAGGAGAACAGGAGTCTCTCCAGCCCTCTTCGGACTTTCAGAAATCGTAGA	994	296	pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa	316	995	CATTAGCATCAACAGGCCCGCGCGTGGCCCGCGGAGAGCACCGCTGGTCACTGT	1054	316	lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe	336	1055	TACCAGGACAGACAACAGATTTTAGAGGCGGAGTTCCCAAGGCTGGCCGAGGCTGTGTC	1114	336	rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy	356	1115	GTTCGTGGCGCTCGTGACGGCTACTTCCGGCTGACCAAGGACTCCAGCACTTCTCTCTG	1174	356	sLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProIle	376	1175	CAAGGAGTGGCACCGCCGAGGCTGTGGAGGAAGTGGCCGAGCAGTGCACCGGCCCAT	1234	376	eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIle	396	1235	CACCTCGACTTTGGCCATCAACAGCTCAAGACTGGGGGCTCACGCTCTGCTCTCTATGT	1294	396	eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr	416	1295	TCCTCCGCGCAGCCCCCAGGACTTTTGACAGTTTCTCTCTACTGTCTGTGTCCAGAACCC	1354	416	oLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerIle	436	1355	CTTTGGTCTCTGATTATAAGGCTGCCTCATCCGCGGACGCCACAGGAACTTCTCTTCT	1414	436	uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTyrAsnSe	456	1415	GGTTGGCCTCAGCCGACCCACAGCATGTCCTTCGAGAGTCTCTGGCAACTGCTGGGATCG	1474	456	rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProly	476	1475	GGGCTGCACGTAGATGGGTGGCAGTACCCCTCACTTCTCTGCTATCCCCAGACCCAA	1534	476	sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy	496
-----	--	-----	--	-----	-----	--	-----	-----	---	-----	-----	---	-----	-----	--	-----	-----	--	-----	-----	--	-----	-----	--	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	---	-----	-----	--	-----	-----	--	-----	-----	--	-----	-----	--	------	-----	--	-----	------	---	------	-----	--	-----	------	---	------	-----	--	-----	------	---	------	-----	---	-----	------	---	------	-----	--	-----	------	--	------	-----	---	-----	------	--	------	-----	---	-----	------	--	------	-----	--	-----	------	---	------	-----	--	-----

Alignment Scores:	0	Length:	3620
Pred. No.:	4597.00	Matches:	895
Score:	86.80%	Conservative:	65
Percent Similarity:	80.92%	Mismatches:	123
Best Local Similarity:	78.45%	Indels:	24
Query Match:	4	Gaps:	7
DB:			

  

US-09-397-967A-16 (1-1099) x US-09-016-434-1055 (1-3620)			
QY	1	MetAlaProSerGluGluThrProIeuIleProGlnArgSerCysSerLeuSer	20
Db	96	ATGGCACTCCAGTGAAGAGACGCCCTGATCCCTCAGCGTTTCATGCAGGCTCTTTGTGCC	155
QY	21	SerGluAlaGlyAlaLeuHisValIeuLeuProArgGlyProGlyProProGlnArg	40
Db	156	ACGAGGCTGTGCTCCATGTGCTGTGCTGCCGCTCGGGGCCCGGGCCCCCAGCGC	215
QY	41	LeuSerPheSerPheGlyAspTryIeuAlaGluAspLeuCysValArgAlaAlaIysAla	60
Db	216	CTATCTTTCTCTTTGGGACCACTTGGCTCAGGACCTGTGCTGCAGGCTGCCAAGGCC	275
QY	61	CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys	80
Db	276	AGCGGCATCCTGCTGTGTACCACTCCCTCTTTGCTCTGGCCACGAGGACCTGCTCTGC	335
QY	81	TrpPheProCysHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr	100
Db	336	TGGTTCCGCCGAGCACATCTTCTCCGTGGAGGATGCCAGCACCAAGTCGTCTGTAC	395
QY	101	ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu	120
Db	396	AGGATTCGCTTTTACTTCCCAATTTGGTTGGGCTGGAGAGTGCACCGCTTCGGGCTA	455
QY	121	ArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln	140

Db 1535 AGAAAGTCAACCTGATCGTGTCCAGAGAGGTCAAGCCCAACATCATCTTGGT 1594  
QY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
Db 1595 TCAGCCCAATCCATACCACTGAGTCAAGTCAATTTCAAGATCCCTGCTGACAG 1654  
QY 515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535  
Db 1655 CTGGAGTGGCATGAGAACCTGGGCGCATGGTCTCTCACCAAGATTACCGGGGCTGTGC 1714  
QY 535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555  
Db 1715 CATGAGTGGTGGATGGGAGGCCCGAAGACAGAGGTGCTCTGAAGGTCAATGATGC 1774  
QY 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575  
Db 1775 CAAGCAAGAAGTCTGATGGATGATCTCTGGAGCAGGAGCTTGATGAGCCAGTGTG 1834  
QY 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG1 595  
Db 1835 GTACCGGCATCTGCTGCTCCACGGCGTGTGATGGTGGAGACAGCACCATGGTGA 1894  
QY 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615  
Db 1895 GGAATTTGTACACCTGGGGGCCATAGACATATATCTGCCAAAACGTGGCCACCTGGTGCC 1954  
QY 615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635  
Db 1955 AGCCAGCTGGAAGTGCAGGTGTCCTGCTCCCGAAGGTGCTCTGGCTCGGGAGGGGC 2074  
QY 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGlyG1 655  
Db 2015 CAAGCGCTGCCCCATGCAATGTCTCTGCCGGAAGGTGCTCTGGCTCGGGAGGGGC 2074  
QY 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675  
Db 2075 TGATGGAGCCCGCCCTTCATCAAGCTGAGTGACCTGGGGTCAGCCCGCTGTGTTAAG 2134  
QY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaG1 695  
Db 2135 CTGGAGATGTCACCGACAGGATCCCTGGGTGGCCCCCGAGTGTCTCCGGAGGGCGCA 2194  
QY 695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheG1 715  
Db 2195 GACACTTAGCTTGAAGTGAACAGTGGGGCTTCGGCGCCAGCGTCTGGGAAGTGTTAG 2254  
QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735  
Db 2255 TGGCGTCACCATGCCATCAGTGCCCTGGATCTGCTTAAGAACTCCAAATTTATAGGA 2314  
QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755  
Db 2315 CCGGACGAGCTGCGGGCCCCCAAGTGGACAGAGTGGCCCTGTGATTCAACAGTGCAT 2374  
QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775  
Db 2375 GGCCTATGAGCCGCTCCAGAGCGCTCTCTCCGAGCGCTCATTCGTGACCTCAATAGCT 2434  
QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795  
Db 2435 CAICTCTTCAGACTATGAGTCTCTCTCAGACCCACACCTGGTCCCTGGCCACCTCGTGA 2494  
QY 795 pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG1 815  
Db 2495 TGGGCTGTGG---AATGGTGGCCAGCTCTATGCTGCGAAGCCACAGATCTTCGAGGA 2551  
QY 815 uArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGGCAACTTTGGCAGCGTGGAGCTGTG 2611  
QY 835 sArgTyrAspProLeuGlyAspAnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855  
Db 2612 CCGCTATGCCCGCTAGGCGCAATACAGGTGGCCCTGTGGTGGCGGTGAACAGCTGCAGCA 2671

QY 855 sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875  
Db 2672 CAGCGGGCCAGACACAGCAGAGGACTTTTCAGCGGAGATTTCAGATCTCTCAAGCACTGCA 2731  
QY 875 sSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895  
Db 2732 CAGTGATTTCAATTGTCAAGTATCGTGTGTACGTATGGCCCGGGCCGAGACCTGGCG 2791  
QY 895 gLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgG1 915  
Db 2792 GCTGTGTCATGGAGTACCTGCCAGCGGCTGCTTGGCGACTTCCTGCAGCGGCACCGCGC 2851  
QY 915 y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTr 934  
Db 2852 GCGCTTCGATGCCAGCGGCTCTCTCTCTATTTCTCGCAGATCTGCAAGGGCATGGAGTA 2911  
QY 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954  
Db 2912 CTGGGCTCCCGCGCTGCTGCGACCGACCTGGCGCCCGCAACATCTCTGTGGAGAG 2971  
QY 954 rGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAs 974  
Db 2972 CGAGGCACACGTCAGATCGCTGACTTCGGCTTAGCTAGCTGCTGCCCTTCACAAAGA 3031  
QY 974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLe 994  
Db 3032 CTACTACGTGGTCCGAGCGAGCGCAGAGCCCATTTTCTGGTATGCCCGGAATCCCT 3091  
QY 994 sSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrG1 1014  
Db 3092 CTCGGACAACTCTTCTTCGCCAGTCAGACGCTCGAGCTTCGGGGTCTGCTGTACGA 3151  
QY 1014 uLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetG1 1034  
Db 3152 GCTCTTCACTTACTCGCACAAAGCTGCGAGCCCTCGCCGAGTTCTTCGGAGTATGATGG 3211  
QY 1034 yProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgAr 1054  
Db 3212 ATGTAGCGGGATGTCCCGCCCTCTGCGGCTCTTGGAACTGTCTGGAGGAGGGCCAGAG 3271  
QY 1054 gLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074  
Db 3272 CTGCGCGGCTCTGCTGCTGCCCTGTGAGGTG----- 3305  
QY 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094  
Db 3306 -----AGCGCGCAGGGGTAGCTCAGTTTCCAGTCTGTAGAT-----TG 3346  
QY 1094 pArgGlyArgProGly 1099  
Db 3347 GGCGGGGCTCTCGGGC 3362

## RESULT 3

US-08-357-598-1  
; Sequence 1, Application US/08357598  
; Patent No. 5705625  
; GENERAL INFORMATION:  
; APPLICANT: Civin, Curt I.  
; APPLICANT: Small, Donald  
; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS



SOFTWARE: PatentIn Release #1.0, Version #1.30

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/357,598

FILING DATE: 15-DEC-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Haile, Lisa A.

REGISTRATION NUMBER: 38,347

REFERENCE/DOCKET NUMBER: 07265/033001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619/678-5070

TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 3807 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-357-598-1

Alignment Scores: 0 Length: 3807  
Pred. No.: 4418.00 Matches: 870  
Score: 86.94% Conservative: 62  
Percent Similarity: 81.16% Mismatches: 130  
Best Local Similarity: 75.39% Indels: 12  
Query Match: 1 Gaps: 6  
DB: 1

US-09-397-967a-16 (1-1099) x US-08-357-598-1 (1-3807)

```
QY 1 MetAlaProSerGluThrProGluLeuProGluLeuSerSerSer 20
Db 168 ATGGCACTCAAGTGAAGACGCGCCCTGATCCCTCAGCGTTCATGACGCCCTTGTGCC 227
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProGlyProGlyProGluArg 40
Db 228 ACGGAGGCTGGTCCCTCCGATGCTGCTGCGCGCTCGGGGCCCGGGGCCCGCCAGCGC 287
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60
Db 288 CTATCTTCTTTGGGACCACTTGGCTGAGGACCTGTGCGGTGACAGGCTGCCAAGGCC 347
QY 61 CysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80
Db 348 AGCGCATCTCGCTGTGTACACTCCCTCTTCTGCGCAGGAGGACCTGCTCTGC 407
QY 81 TrpPheProProSerHisLeuPheCysIleGluAspValAspThr-GlnValLeuValTy 100
Db 408 TGGTT-CCCCGAGCCACATCTTCCGTGGAGGATGCCAGCACCCCAAGTCTGTGTA 466
QY 100 rArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLe 120
Db 467 CAGGATTCGCTTTTACTTCCCAATTTGTTGGGCTGGAGAAGTGCCACCGCTTCGGGCT 526
QY 120 uArgLysAspLeuThrSerAlaLeuAspLeuHisValLeuGluHisLeuPheAlaG1 140
Db 527 ACGCAAGGATTTGGCCAGTGTATCTTACCTGACCTGCGCAGTCTGGAGACCTTTTGCCCA 586
QY 140 nHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnG1 160
Db 587 GCACCGCAGTACCTGGTGTAGTGGCGCTCCCGTGGCTCAGTCTCAAGGAGCAGG 646
QY 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnAr 180
Db 647 TGAGTGTCTCAGCCTGGCGGTGTTGACCTGCGCCGATGCGCGAGAGCAGGCCAGCG 706
QY 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200
Db 707 GCGGGGAGGCTGCTGAAGACGTGTGAGCTACAGGCGCTTGCCTACCCCAAGCCTCGCGA 766
QY 200 pValIleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAl 220
Db 1843 CCAAGCACAGAACTGTCATGGAGTCTATTCCTCGAAGCAGCGAGCTTGTATGACCAAGTGT 1902
```

```
Db 767 CCGTATCCAGGCGCTGAGCTTCGTACCGGGAGGCGGTATTCGAGGACGCTGAGAGCCCC 826
QY 220 aLeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleL 240
Db 827 CCGTGG-CCGGTGGCCGCTGCCAGGACGACCGGACCTCGCTCATGGCAAAGTACATCA 885
QY 240 euAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260
Db 886 TGGACCTGGAGCGGCTGGATCCAGCCGGGCGCGGAGACCTTCCAGTGGGCGCTCCCTG 945
QY 260 lyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleP 277
Db 946 GGGCCCTTGGTGGCCACGAGCGGCTGGGCTCGTCCGCTGCTGCTGAGCGCGGATCG 1005
QY 277 roTripSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValA 296
Db 1006 CTTGGACCCAGGAGAACAGAGGCTCTCCAGCCCTTCTGCGACTTTCAGAAATCGTAG 1065
QY 296 spValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrV 316
Db 1066 ACATTAGCATCAAGCAGGCGCGCGGCTTGGCCCGGCGAGAGCACCGCTGCTCACTG 1125
QY 316 alThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuS 336
Db 1126 TTACCAGGACAGAACACAGATTTTAGAGCCGAGTTCCAGGCGCTGCCAGGCTCTGT 1185
QY 336 exPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheC 356
Db 1186 CGTTCGTGGCGCTCGTGGAGCGCTACTTCCGGCTGACACCGGACTCCAGCACTTCTTCT 1245
QY 356 yLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376
Db 1246 GCAAGGAGGTG---GACCCGAGGCTGCTGGAGAAAGTGGCGGAGAGTGCACGCGCCCA 1302
QY 376 leThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396
Db 1303 TCACCTCGGACTTTGCCATCACAGCTCAAGACTGGGGCTCACGCTCTGCTGCTCTATG 1362
QY 396 leLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrP 416
Db 1363 TTCTCCGCCCATCCCCAGACTTTTGACAGCTTCTCTCTCACTGTGTGTCTCAGAAC 1422
QY 416 roLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerL 436
Db 1423 CCCTTGTCTGTATTAAAGGCTGCTCATCCGGCGCAGCCCGCAGGAAACCTTCTCTTC 1482
QY 436 euValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn 456
Db 1483 TGGTGTGGCTCAGCGGACCCCGCAGAGCTCTTCGAGAGCTCTTGCAACCTGCTGGATG 1542
QY 456 erGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProL 476
Db 1543 GGGGGCTGCAGTAGATGGGTGGCAGTGCACCTTCTCTGCTGTATCCCGACACCA 1602
QY 476 ysGluLysSerAsnLeuIleValValArgGlyCysAsnProAlaProAlaProGlyC 496
Db 1603 AAGAAAGTCCAACTGATTGTGTGTGTCAGAGAGTTCACAGCCCGCAGCCCATCTCTTGG 1662
QY 496 ysSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAsp 515
Db 1663 TTCAGCCCCAATCCCAATACAGCTGTAGTTCAGTACATTTTCAAGATCCCTGCTGACA 1722
QY 515 erLeuGluTyrPheIleGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerA 535
Db 1723 GCCTGGAGTGGCATGAGAACCTGGGCGCATGGCTCTTCCACCAAGATTTACCGGGGCTGC 1782
QY 535 rGArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAsp 555
Db 1783 GCCATGAGTGGTGGATGGGAGGCCCGAAAGACAGAGAGGTGCTGCTGAAGTCACTGATG 1842
QY 555 erArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValS 575
Db 1843 CCAAGCACAGAACTGTCATGGAGTCTATTCCTCGAAGCAGCGAGCTTGTATGACCAAGTGT 1902
```



QY 575 erTyrProHisLeuValLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595  
 Db 1903 CGTACGGCATCTCGTCTCTCCACGGCGTGTGATGGTGGAGACAGACCATGGTGC 1962  
 QY 595 lngluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValS 615  
 Db 1963 AGGAATTTGTACACCTGGGGGGCATAGACATGATCTTCGAAACCGTGGCCACCTGGTGC 2022  
 QY 615 erAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluA 635  
 Db 2023 CAGCCAGCTGGAGCTGAGGTGGTCAACAGCTGGCTTACGCCCTCAACTATCTGGAGG 2082  
 QY 635 spLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyG 655  
 Db 2083 ACAAAAGCCCTGTCCTCCATGGCAATGTCTCTGCCCGGAAGGTGCTCTCGGTGGGAGGGG 2142  
 QY 655 lYAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuS 675  
 Db 2143 CTGATGGGAGCGCCCTTCATCAAGCTGAGTGACCTGGGGTTCAGCCCGCTGTGTAA 2202  
 QY 675 erLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlag 695  
 Db 2203 GCCTGGAGATGCTACCGACAGATCCCTGGGTGGCCCCCGAGTGTCTCCGGAGGGCG 2262  
 QY 695 lngluCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheG 715  
 Db 2263 AGACACTTAGCTTGAAGCTGACAGTGGGGCTTCGGCGCCACGGTCTGGGAAGTGTATA 2322  
 QY 715 lngArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluA 735  
 Db 2323 GTGGCGTCCACCATGCCCTAGTTCAGTGGCTTCTGCTTGAAGAACTCCAAATTTATGAGG 2382  
 QY 735 spGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysM 755  
 Db 2383 ACCGCGCAGCTGTCGGCCCCCAAGTGGACAGAGCTGGCCCTGCTGATTCACACAGTGA 2442  
 QY 755 etAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775  
 Db 2443 TGGCTATGAGCGGTCCAGAGGCGCTCTCTACGAGCGTCTATCGTACCTCATAGTC 2502  
 QY 775 euIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgA 795  
 Db 2503 TCATCTCTTACAGTATGAGTCTCTCTCAGACACACC---TGGTCCCTGGCACTCGTG 2559  
 QY 795 spGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG 815  
 Db 2560 ATGGCTGTGG---AATGGTCCCGAGCTCTATGCTGCCAGACCCCGATCTTCGAGG 2616  
 QY 815 luArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuC 835  
 Db 2617 AGAGACACCTCAAGTACATCTCAGAGCTGGGCAAGGGCTTCTTTGGCAGCGTGGAGCTGT 2676  
 QY 835 ysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnH 855  
 Db 2677 GCCGCTATGACCGGTAGCGACATACAGGTGCCCTGGTGGCGGTGAACAGCTGCAGC 2736  
 QY 855 isSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLysAlaLeuH 875  
 Db 2737 ACAGCGGGCCAGACAGCAGAGGACCTTCAGCGGGAGATTTCAGATCTCTCAAGACAGC 2796  
 QY 875 isSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuA 895  
 Db 2797 ACAGTATTTTCATGTCAAGTATCGTGTGTGTCAGTATGGCCCGGCCGACAGACCTGT 2856  
 QY 895 rgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgG 915  
 Db 2857 CGCTGGTATGAGTACTGTCGCCAGCGCTCTTGGCGCATCTTCTGCGCGGCAACCGG 2916  
 QY 915 lYLeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrL 935  
 Db 2917 GCCTCGATGCCAGCGGCTCTTCTTATTCTTCGTCAGATCTGCAAGGGCATGGAGTACC 2976

QY 935 euGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerG 955  
 Db 2977 TGGGTCTCCCGCCCTCGTGCACCGGACCTGGCCGCCGAAACATCTCGTGGAGAGCG 3036  
 QY 955 luAlaHisValLysIleAlaAspPheGlyLeuAlaLysIleLeuProLeuGlyLysAspT 975  
 Db 3037 AGSCACACGTCAGATCGCTGACTTGGCCCTAGCTAAGCTGCTGCCCTTGACAAAGACT 3096  
 QY 975 yTrpValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuS 995  
 Db 3097 ACTACGTGTGGCGAGCCAGGCGCAGCCCAATTTCTGGTATGCCCGCAATCCCTCT 3156  
 QY 995 erAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluL 1015  
 Db 3157 CGACAAACATCTTCTCTCGCAGTCAGAGCTTGGAGCTTGGGGTCTGCTGTACGAGC 3216  
 QY 1015 euPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyP 1035  
 Db 3217 TCTTCACTTACTGCGACAAAGCTGCGAGCCCTCGGCCGAGTTCTTGGGATGATGGAT 3276  
 QY 1035 roGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArgL 1055  
 Db 3277 GTGAGCGGATGTCCCGCCCTCTGCGGCTCTTGGAACTGCTGGAGGAGGCCAGAGC 3336  
 QY 1055 euProProProProThrCysProThrGluVal 1065  
 Db 3337 TGCCGGCGCTCTTGTGCTGCCCTGTGAGGTG 3368

## RESULT 4

US-09-003-289-1  
 ; Sequence 1, Application US/09003289  
 ; Patent No. 5916792  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Civin, Curt I.  
 ; APPLICANT: Small, Donald  
 ; TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Fish & Richardson P.C.  
 ; STREET: 4225 Executive Square, Suite 1400  
 ; CITY: La Jolla  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 92037  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/003,289  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/357,598  
 ; FILING DATE: 15-DEC-1994  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Haile, Lisa A.  
 ; REGISTRATION NUMBER: 38,347  
 ; REFERENCE/DOCKET NUMBER: 07265/033001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 619/678-5070  
 ; TELEFAX: 619/678-5099  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3807 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA  
 ; US-09-003-289-1

Alignment Scores:

Pred. No.:	0	Length:	3807
Score:	4418.00	Matches:	870
Percent Similarity:	86.94%	Conservative:	62
Best Local Similarity:	81.16%	Mismatches:	130
Query Match:	75.39%	Indels:	12
DB:	2	Gaps:	6

US-09-397-967A-16 (1-1099) x US-09-003-289-1 (1-3807)

QY	1	MetAlaProSerGluThrProLeuIleProGlnArgSerCysSerIeuSerSer	20
Db	168	ATGCACCTCCAAAGTGAAGAGACGCCCTGTATCCCTCAGCGTTTCATGAGCCTCTTGTC	227
QY	21	SerGluAlaGlyAlaLeuHisValIeuLeuProProArgGlyProGlyProGlnArg	40
Db	228	ACGGAGGCTGTGCTGCTGATGCTGCTGCCCGCTCGGGGCCCGGGCCCCCAGCGC	287
QY	41	LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValAlaAlaLysAla	60
Db	288	CATATCTTCCTTTGGGAGCACCTGGCTGAGGACCTGTGCGTGCAAGGCC	347
QY	61	CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys	80
Db	348	AGCGCATCCTGCTGTGTACCACTCCCTTTTGCTTGCCACGGAGACTGTCCTGC	407
QY	81	TripPheProSerHisIlePheCysIleGluAspValAspThr-GlnValLeuValTy	100
Db	408	TGGTT-CCCCGAGCCACATCTTCTCGTGGAGATGCCAGACCACCCAGTCTCTGTATA	466
QY	100	rArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLe	120
Db	467	CAGATTCTGCTTTACTTCCCATTGTTGGCTGGAGAAGTGCACCGCTTGGGCT	526
QY	120	uArgLyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisIeuPheAlaGl	140
Db	527	ACGCAAGATTGGCCAGTGTATCTCTGACCTGCCAGTCTCGGAGCCTCTTTCGCCA	586
QY	140	nHisrSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIySgluIncl	160
Db	597	GCACCGTAGTACCTGTGTAGTGGCGCTCCCCGTGGCTCAGICTCAAGGAGCAGGG	646
QY	160	yGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluInlalnAr	180
Db	647	TGAGTGTCTCAGCTGGCGGTGTGGACCTGGCCGGATGGCGAGAGCAGGCCAGCG	706
QY	180	gProGlyGluIeuLeuYsThrValSerTyrLysAlaCysLeuProSerLeuArgAs	200
Db	707	CGGGGAGAGCTGTCTGAAGACTGTCTCAGCTACAGGCTGTCTACCCCAAGCTCGCGA	766
QY	200	pValIleGlnGlyGlnAsnPhelThrArgArgArgileArgArgThrValValLeuAl	220
Db	767	CTGTATCCAGGCCCTGAGCTTCGTGACGGGAGCGGTATTCGAGGACGGTGGAGAGCCC	826
QY	220	aLeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIystrIleL	240
Db	827	CTGTGG-CCGGGTGCGCTGCTGCCAGGAGACCGGCACTGCTCATGGCAAGTACATCA	885
QY	240	euaSpLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG	260
Db	886	TGAGCTGTGAGCGGTGGATCCAGCGGGGCCCGCGAGACCTTCCACGTGGCGCTCCCTG	945
QY	260	lyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleP	277
Db	946	GGGCCCTTGTGTGGCCACGACCGGGCTGGGCTGTCTCGCGTGGCTGGTACGGCGCATCG	1005
QY	277	roTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleVala	296
Db	1006	CTGTGACCCAGGAGAACAGGAGGTCTCTCCAGCCCTTCTGCGACTTCCAGAAATCTGAG	1065
QY	296	spValSerIleAsnGlnAlaProArgValGlyProAlaGlyIleGluHisArgLeuValThrV	316
Db	1066	ACATTAGCATCAGACGAGGCCCGCGGTGTGGCGCGCGGAGAGCAGCCGCTGTGTCACTG	1125

Db 2203 GCTGGAGATGTCACCGACAGATCCCTGGTGGCCCGGAGTGTCTCCGGAGGCGC 2262  
Qy 695 lnThrLeuCysLeuGluAlaAepLysTrpGlyPheGlyAlaThrThrTrpGluValPheG 715  
Db 2263 AGACACTTAGCTTGGAAAGTGTACAAGTGGGGCTTCGGCGCCACGGTCTGGAAAGTGTGA 2322  
Qy 715 lnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluA 735  
Db 2323 GTGGGGTCAACCATGCCCATCAGTGCCTAGATCCTCTGAAGAACTCCCAATTTATGAGG 2382  
Qy 735 spGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysM 755  
Db 2383 ACCGGCAGAGTGTTCGGCCCGCCCAAGTGGACAGAGCTGGCCCTGCTGATTCACAGTGCA 2442  
Qy 755 etAlaTrpAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyI 775  
Db 2443 TGGCCATGAGCGGGTTCAGAGGCCCTCTTAGAGCCGTGATCTGTGACCTCAATAGTC 2502  
Qy 775 euIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgA 795  
Db 2503 TCATCTCTTACACTATGAGTCTCTCTCAGACCACACC---TGGTGCCCTGGCAGCTCGTG 2559  
Qy 795 spGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG 815  
Db 2560 ATGGGCTGTGG---AATGGTGGCCAGCTCTATGCTGCCAAGACCCACGATCTTCGAGG 2616  
Qy 815 luArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuC 835  
Db 2617 AGAGACACTCAAGTACATCTCACAGCTGGGCAAGGGCTTCTTGGACGCTGGAGCTGT 2676  
Qy 835 ysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnH 855  
Db 2677 GCCGCTATGACCCGTAGCGCAATACAGGTGCGCTGGTGGCCGTGAAGACAGCTGCAGC 2736  
Qy 855 isSerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuH 875  
Db 2737 ACAGGGCCACAGCAGCAGAGGAGCTTTCAGCGGGAGATTGATCTCTCAAGCACAGC 2796  
Qy 875 isSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuA 895  
Db 2797 ACAGTATTTATGTCAAGTATCTGTGTGTGATGCTATGGCGCCGGCCGAGACCCCTG 2856  
Qy 895 rgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg 915  
Db 2857 CGCTGGTCAAGTACCTGCCAGCGCTGCTGGCGGACTTCTTGCGAGCGGACCGGG 2916  
Qy 915 lyLeuHisThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrL 935  
Db 2917 GCCTCGATGCCAGCGCGCTCTCTCTATCTCGCAGATCTGCAAGGGCATGGAGTACC 2976  
Qy 935 euGlyAlaArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerG 955  
Db 2977 TGGGCTCCCGCGCTGTGTGCACCGACCTGGCGCCCGCCGAAACATCTCTGTGAGAGCG 3036  
Qy 955 luAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspT 975  
Db 3037 AGGCACACGTCAAGATCGTCACTTGGCCCTAGCTAGCTCTGCGGCTTGACAAAGACT 3096  
Qy 975 yTrpValValArgGluProGlyGlnSerProIlePheTyrTrpAlaProGluSerLeuS 995  
Db 3097 ACTACGTGGTCCGAGCCAGCCAGAGCCCATTTCTGTGTATGCCCCCGAATCCCTCT 3156  
Qy 995 erAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluL 1015  
Db 3157 CGGACACATCTCTCTCGCAGTCAAGAGTCTGGAGCTTCGGGGTCTGCTGTACGAGC 3216  
Qy 1015 euPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyP 1035  
Db 3217 TCTTCACCTACTGCGACAAAGCTGCAGCCCTCGGCCGAGTTCCTGCGGATGATGGAT 3276  
Qy 1035 roGluArgGluGlyProProLeuCysArgLeuLeuLeuLeuAlaGluGlyArgArgL 1055

Db 3277 GTGAGCGGATGTCCCGCCCTCTGGCGCTCTTGAACCTGTGGAGAGGCCAGAGGC 3336  
Qy 1055 euProProProProProProProProProProProProProProProProProProPro 1065  
Db 3337 TCCCGCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3368  
RESULT 5  
PCT-US95-16435-1  
Sequence 1, Application PC/TUS9516435  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University School of Medicine  
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16435  
FILING DATE: 15-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033W01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
PCT-US95-16435-1  
Alignment Scores:  
Pred. No.: 0 Length: 3807  
Score: 4418.00 Matches: 870  
Percent Similarity: 86.94% Conservative: 62  
Best Local Similarity: 81.16% Mismatches: 130  
Query Match: 75.39% Indels: 12  
DB: 5 Gaps: 6  
US-09-397-967a-16 (1-1099) x PCT-US95-16435-1 (1-3807)  
Qy 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
Db 168 ATGCACCTCCAAAGTGAAGAGACGCCCTGATCCCTCAGCGTTTCATGACGCTCTTGTTC 227  
Qy 21 SerGluAlaGlyAlaLeuHisValLeuProProArgGlyProGlyProGluArg 40  
Db 228 ACGAGAGCTGTGCGCTTCATGTGCTGCTGCGGCGCCCGGCGCCCGCCAGCGC 287  
Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
Db 288 CTAATCTTCTCTTGGGGACCACTTGGCTGAGGACCTGTGCTGAGGCTGCCAAGGCC 347  
Qy 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
Db 348 AGCGCATCTGCTGTGTACCACTCCCTCTTGTCTGTGCGCCACGAGGACCTGCTCTGC 407  
Qy 81 TrpPheProProSerHisIlePheCysIleGluAspValAspThr-GlnValLeuValTy 100

Db 408 TGGTT-CCCCGAGCCACATCTTCTCCGTGGAGGATGCCAGCACCCCAAGTCTCTGTGTA 466  
 Qy 100 rArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLe 120  
 Db 467 CAGGATTGCTTTACTTCCCAATTGGTTGGCTGGAGAGTGCCACCGCTTGGGGCT 526  
 Qy 120 uArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaG 140  
 Db 527 AGCGAAGGATTTGGCCAGGCTATCTTCACTGCGCACTCTGGAGCACCTTTTGCCCA 586  
 Qy 140 nhisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln 160  
 Db 587 GCACCGCATGACCTGGTGGTGGGGCGCTCCCGCTGGCTCAGTCTCAAGAGCAGGG 646  
 Qy 160 yGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnAr 180  
 Db 647 TGAGTGTCTCAGCCTGGCGTGTGTGACCTGGCCGGATGGCGGAGAGCAGGCCAGCG 706  
 Qy 180 gProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAs 200  
 Db 707 GCGGGAGAGCTGCTGAAGACTGTCACTCAAGGCTGCTTACCCCAAGCTCGCGCA 766  
 Qy 200 pValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAl 220  
 Db 767 CTGTATCCAGGGCTGAGCTTCGTGACGGGGGCGGTATTTCGGAGGACGGTGGAGAGCCC 826  
 Qy 220 aLeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleL 240  
 Db 827 CTGTGCG-CCGGTGGCGCTGCTGACGACAGCAGCGGCTGCTGCTGCTGCTGCTGCTG 885  
 Qy 240 euAspLeuGluArgLeuHisPheAlaAlaThrThrGluThrPheArgValGlyLeuProG 260  
 Db 886 TGACCTGGAGCGCTGGATCCAGCGGGCGCGGAGACTTCCACGTGGGCTCCCTG 945  
 Qy 260 lYAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleP 277  
 Db 946 GGGCCCTTGGTGGCCACAGCGGGTGGGCTGCTGCGCTGGTGGTGGTGGTGGTGGTGG 1005  
 Qy 277 roTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValA 296  
 Db 1006 CTTGGACCCAGGAGAACAGAGGCTCTCCAGCCCTTCTCGACTTTTCCAGAAATCGTAG 1065  
 Qy 296 spValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrV 316  
 Db 1066 ACATTAGCATCAACAGCGCCCGCGCTTGGCCCGCGGAGACACCGCTGTCTACTG 1125  
 Qy 316 alThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuS 336  
 Db 1126 TTACCGAGACAGACCAACAGATTTTAGAGGCGGAGTTCCACAGGCTGCGCGAGGCTCTGT 1185  
 Qy 336 erPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheC 356  
 Db 1186 CGTTCGTGGCGCTCGTGACGCGCTACTTCCGCGTGACCGAGCTCCGAGCACTTCTTCT 1245  
 Qy 356 yLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376  
 Db 1246 GCAGGAGGAGTG---GACCCGAGGCTGCTGGAGGAGTGGCGAGAGTCCACGCGCCCA 1302  
 Qy 376 leThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396  
 Db 1303 TCACCTGGACTTTGGCCATCAACAGCTCAAGACTGGGGGCTCAGTCTCGCTCTCTCTATG 1362  
 Qy 396 leLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrP 416  
 Db 1363 TTCTCCGCGCATCCCGAGGACTTTGACGCTTCTCCTCCTCCTGCTGTGTCCAGAAC 1422  
 Qy 416 roLeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerL 436  
 Db 1423 CCCTTGGTCTGATTATTAAGGCTGCTCCATCCGCGGAGCCGCCACAGAACTTCTCTTC 1482  
 Qy 436 euValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaCysTrpAsnS 456  
 Db 1483 TGGTTGGCTCAGCGCACCCACAGAGTCTTTCGAGAGCTCTCTGGCAACCTGCTGGGATG 1542

Qy 456 erGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProL 476  
 Db 1543 GGGGGCTGCACGTAGATGGGTGGCAGTGCACCTCACTTCTCTGTATATCCCGACACCA 1602  
 Qy 476 yGlyLysSerAsnLeuIleValArgArgGlyCysAsnProAlaProAlaProGlyC 496  
 Db 1603 AAGAAAAGTCCAACTGATTGGTGGTCCAGAGAGGTACAGCCCAACCCACATCATCTTGG 1662  
 Qy 496 ySerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspS 515  
 Db 1663 TTCAGCCCCAATCCCAATACAGCTGAGTCAGATGACATTTTCAAGATCCCTGCTGACA 1722  
 Qy 515 erLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerA 535  
 Db 1723 GCTGGAGTGGCATGAGAACCTGGGCGCATGGTCTTCCCAAGATTTTACCGGGGCTGTC 1782  
 Qy 535 rArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspS 555  
 Db 1783 GGCATGAGTGGTGGATGGGAGGCGCCGAAAGACAGAGGTGCTGCTGAAGGTTCATGATG 1842  
 Qy 555 erArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValS 575  
 Db 1843 CCAAGCAAGAACTGCATGGAGTCACTCTGGAAGCAGCGAGCTTGTATGAGCAAGTGT 1902  
 Qy 575 erTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValG 595  
 Db 1903 CGTACCGCATCTCTGCTGCTCCAGCGGTGTGTCATGGTGGAGACAGCACCATGTCTG 1962  
 Qy 595 lncLupheValTyrLeuGlyAlaIleAsnMetTyrLeuArgLysArgGlyHisLeuValS 615  
 Db 1963 AGGAATTTGACACTGGGGGCGCATAGACATGATATCTGCCAAACGCTGGCCACTGTGTC 2022  
 Qy 615 erAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluA 635  
 Db 2023 CAGCCAGCTGGAAGCTGCAGTGGTCAACAGCTGGCTTACGCCCTCAACTATCTGAGG 2082  
 Qy 635 spLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyG 655  
 Db 2083 ACAAGGCTGTCCCATGGCAATGCTCTGCGCGAAGAGTGTCTCTGCGTGGGAGGGGG 2142  
 Qy 655 lyAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuS 675  
 Db 2143 CTGATGGAGCCCGCCCTTATCAAGCTGAGTGAACCTGGGTGAGCCCGCTGTGTATA 2202  
 Qy 675 erLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaG 695  
 Db 2203 GCCTGGAGATGCTCACCGACAGGATCCCTGGGTGGCCCGAGTGTCTCGGGAGGCGC 2262  
 Qy 695 lncThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheG 715  
 Db 2263 AGACACTTAGCTTGAAGCTGCAAGTGGGCTTTCGCGCCACCGGTCTTGGGAAGTGTATA 2322  
 Qy 715 lncArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluA 735  
 Db 2323 GTGGCGTCACCATGCCCATCAGTCCCTAGATCTCTGTAAAGAACTCCAATTTATGAGG 2382  
 Qy 735 spGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysM 755  
 Db 2383 ACCGGCAGCAGCTGTCGGCCCCCAAGTGGACAGAGTGGCCCTGCTGATTCACAGTGA 2442  
 Qy 755 etAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyL 775  
 Db 2443 TGGCTATGAGCCGCTCCAGAGGCGCTCTTACGAGCGCTCATTCGTGACCTCAATATGC 2502  
 Qy 775 eulleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgA 795  
 Db 2503 TCATCTCTTACAGATGAGTCTCTCTCAGACCCACACC---TGGTGGCTTGGCACTCGTG 2559  
 Qy 795 spGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG 815  
 Db 2560 ATGGGCTGTGG---AATGGTGGCCAGCTCTATGCTGCGCAAGACCCACGATCTTCGAGG 2616

815 luArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuC 835  
2617 AGAGACACCTCAAGTATCATCTCAGCTGGGCAAGGCTTCTTTGGCAGCGTGAGCTGT 2676  
835 ysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuH 855  
2677 GCGCTATGACCGCTAGGCGCAATACAGTGCCCTGGTGGCGGTGAACACGCTGCAGC 2736  
855 isSerValProAspGlnGlnArgAspPheGlnArgGluLleGlnLleLeuLysAlaLeuH 875  
2737 ACAGCGGGCGAGACACGAGGAGCTTTTCAGCGGGAGATTCAGATCCTCAAGACAGC 2796  
875 isSerAspPheLleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuA 895  
2797 ACAGTGATTTCAATGTCAAGTATCGTGGTGCTAGCTATGCGCGGCGCGCGAGCCCTG 2856  
895 rglLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgG 915  
2857 CGCTGCTATGAGTACCTGCCAGCGGCTGCTTGCAGCTTTCCTGCAGCGGCACCGG 2916  
915 lyLeuHisThrAspArgLeuLeuPheAlaTrpGlnLleCysLysGlyMetGluTyrL 935  
2917 GCCTCGATGCCAGCGCTCTCTCTATTCCTCGCAGATCTGCAAGGCGCATGGAGTACC 2976  
935 euGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnLleLeuValGluSerG 955  
2977 TGGGTCCCGCGCTGCGTGCACCGCGACTGCGCGCGCGAAACATCTCTGCTGGAGAGCG 3036  
955 luAlaHisValLysLleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspT 975  
3037 AGGCACAGTCAGATCGCTGCTTCCGCTAGCTAAGCTGCTGCGCGCTTGCAAGACT 3096  
975 yTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuS 995  
3097 ACTAGTGTGTCGCGAGCGAGCGAGCGCCCATTTCTGATGATGCCCGCGATCCCTCT 3156  
995 erAspAsnLlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluL 1015  
3157 CGGACAACTCTCTCTCGCAGTCAGAGCTGCGGCTGCGGCTGCTGCTGTCAGAGC 3216  
1015 euPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyP 1035  
3217 TCTTCACTACTGCGCAAAAGCTGCGCCCTCGCGCGAGTCTCGCGGATGATGGGAT 3276  
1035 roGluArgGluGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgArgL 1055  
3277 GTGAGCGGGATGTCGCCCGCTCTGCCGCTCTTGGAACTGCTGAGAGGGGCCAGAGGC 3336  
1055 euProProProThrCysProThrGluVal 1065  
3337 TGGCGGCGCTCTGCTGCGCTGCTGAGGTG 3368

## RESULT 6

US-09-046-158A-21  
; Sequence 21, Application US/09046158A  
; Patent No. 6187552  
; GENERAL INFORMATION:  
; APPLICANT: Robert, Steven L.  
; APPLICANT: Kaytes, Paul S.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF  
; TITLE OF INVENTION: JAK2/CYTOKINE RECEPTOR BINDING  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property  
; STREET: 301 Henrietta Street  
; CITY: Kalamazoo  
; STATE: MI  
; COUNTRY: USA  
; ZIP: 49001  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,158A

FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Darnley Jr., James D.

REGISTRATION NUMBER: 33,673  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 616/833-2210

TELEFAX: 616/833-8897  
TELEX: 224401

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:

LENGTH: 3435 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: cDNA  
US-09-046-158A-21

## Alignment Scores:

Pred. No.: 1.2e-234 Length: 3435  
Score: 2629.00 Matches: 534  
Percent Similarity: 66.21% Conservative: 193

Best Local Similarity: 48.63% Mismatches: 340  
Query Match: 44.86% Indels: 31  
DB: 3 Gaps: 14

US-09-397-967A-16 (1-1099) x US-09-046-158A-21 (1-3435)

QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProG1 39  
Db 111 TCCAGTTCTTCAGGT-----GTATCTTTACCATTCCTTGGAAATCTGAGGCAGA 161  
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaLay 59  
Db 162 TTATCTGACCTTTCATCTGGGAGTATGTTCGAGAGAAATCTGATTGCTGCTCTAA 221  
QY 59 sAlaCysGlyLleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79  
Db 222 AGCTTGCTGATATCACCTGTGATCATATATGTTTCTTTAATGATGAACAGAAAG 281  
QY 79 rCysTrpPheProSerHisLlePheCysIleGluAspValAspThrGlnValLeuVa 99  
Db 282 GATCTGTTATCCACCAACCATGCTTCCATATAGATGAGTCAACAGGCATATGTACT 341  
QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117  
Db 342 CTACAGAAATAAGATTTTACTTTCTCTGTTGTTTATGAGTGGCAGCAACAGAGCTATCG 401  
QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaLleLeuAspLeuHisValLeuGluHisLe 137  
Db 402 GCATGGAATATCTCGAGGTGCTGAGCTCTCTTCTTGTATGACTTTGTCATGCTTACCT 461  
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157  
Db 462 CTTTGTCTCAGTGGCGCATGATTTTGTGTCATGATGATGATGATGATGATGATGATGATG 521  
QY 157 sGluGlnGlyPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177  
Db 522 AACACAGGAAGAATGCTCTGGGATGGCAGTGTGTAGATATGATGATGATGATGATGATGATG 581  
QY 177 nAlaGlnArgProGlyGluLeuLysThrValSerTyrLysAlaCysLeuProProSe 197  
Db 582 CGATCAAAACCCACTGCGCACTATTAACCTATATGATGATGATGATGATGATGATGATG 641  
QY 197 rLeuArgAspValLleGlnGlyGlnAsnPheValThrArgArgArgLleArgThrVa 217  
Db 642 TATTCGAGCAAGATCCAGACTATCATATTTTGGACAGGAAGGAATAGGTACAGATT 701  
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlay 237

1779	GTCTTTCTTTGAACGACAGCATATGATGACCAAGCTTTCTCTCAACAGCATTTGGTTTAA	1838
582	uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuG	601
1839	TTATGGAGTATGTCGTGTGGAGACGAGAATAATTTCTGGTTTCAGGAGTTTGTAAATTTGG	1898
601	yAlaIleAspMetTyrLeuArgIysArgGlyHISLeuValSerAlaSerTrpIysLeuG	621
1899	ATCACTAGATACATATCTGAAAAAGAAATAAAATTTGTATAATATATATATGGAACATTGA	1958
621	nValThrIysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspIysGlyLeuProHisG	641
1959	AGTTGCTAAACAGTTGGCATGGCCATGCAATTTCTAGAGAGAAACACCCCTTATTCATGG	2018
641	yAsnValSerAlaArgIysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnPr	659
2019	GAATGATGTGCCAAAATAATTTCTGCTTATCAGAGAGAAAGACAGAGAGACAGGAAATCC	2078
659	pProPheIleIysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLe	679
2079	TCCTTTTCATCAAACTTAGTGATCTCGCATTTAGTATTACAGTTTGGCCAAAGGACATCT	2138
679	uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe	699
2139	TCAGAGAGATACCATGGGTACCACCTGAATGCATTTGAAATCTCTAAATAATTTAAATTT	2198
699	uGluAlaAspIysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl	719
2199	GGCAACAGACAAATGGAGTTTGGTACCACCTTTGTGGGAAATCTTCGCTGGAGGAGATAA	2258
719	ahisIleThrSerLeuGluProAlaIysIysLeuIysPheTyrGluAspGlnGlyGlnLe	739
2259	ACCTCTAAGTGTCTGGATTTCTCAAGAGAAAGCTACAAATTTATGAGATAGGCATCAGCT	2318
739	uProAlaLeuIysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr	759
2319	TCCTGCACCAAGTGGCGAGATTAGCAAACTTTATAATAATTTATGGATTATGAACC	2378
759	oGlyArgArgProSerPheArgAlaIleLeuArgAspIleuAsnGlyLeuIleThrSerAs	779
2379	AGATTTTCAGGCTTCTTTCAGAGCCATCATACGAGATCTTACAGATTTGTTTACTCCAGA	2438
779	pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy	798
2439	TTATGAACATATTAAACAGAA---AATGACATGTTTACCAATATGAGATAGGTCCTTAGG	2495
798	sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLe	818
2496	GTITTTCTGTGTCG---TTTGAAGACCCGGATCTCACACAGTTTGAAGAGAGACATTT	2549
818	uIysTyrIleSerLeuLeuGlyIysGlyAsnPheGlySerValGluLeuCysArgTyrAs	838
2550	GAATTTCTACAGCAACTTGGCAAGGTTAATTTGGAGGTGGAGATGTCGCCGATATGA	2609
838	pProLeuGlyAspAsnThrGlyProLeuValAlaValIysGlnLeuGlnHisSerValPr	858
2610	CCCTCTACAGGACCAACTGGGAGGTGGTCTGTGTAAAAAAGCTTCAGCATAGTACTGA	2669
858	oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuIysAlaLeuHisSerAspPh	878
2670	AGAGCACCTTAAGAGACTTTGAAAGGAAATTTGAATCTCTGAAATCTCCATCAGCATGACA	2729
878	eIleValIysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMe	898
2730	CATTGTAAGTACAAAGGAGTGTCTACAGTGTGGTGGCGTAATCTAAATAATTAATAT	2789
898	tGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHi	917
2790	GGAAATATTACCATATGGAAGTTTACGAGACTATCTTCAAAAACATAAAGACCGGTAGA	2849
917	sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysIysGlyMetGluTyrLeuGlyAl	937
2850	TCACATAAACTTCTGCAAGTACATCTCAGATATCAAGGGTATGGAGTATCTTGTGATC	2909









Db 3426 CAAAGTTTCTGGGCTCAGATGTTGGAGCTTGGAGTGGTCTGTATGAATTTTCAC 3485  
QY 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr 1037  
Db 3486 ATACATTGAGAGAGTAAAGTCCACCAGGGNAATTTATGCGTATGATTGGCAATGACAA 3545  
QY 1037 gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr 1056  
Db 3546 ACAAGGACAGATGATCGTCTTCATTTGATAGAACTTTTGAAGAATAATGGAAGATTACC 3605  
QY 1056 oProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl 1076  
Db 3606 AAGACACAGATGATGCCAGATGAGATCTATATGATCATCAGAGATGCTGGAACAATAA 3665  
QY 1076 uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093  
Db 3666 TGTAATCAAGCCCTCTCTTAGGATCTAGCTCTTCAGTGGATCAANAATA 3717

## RESULT 8

US-09-196-480-1  
; Sequence 1, Application US/09196480  
; Patent No. 6019966  
; GENERAL INFORMATION:  
; APPLICANT: Coleman, Roger  
; APPLICANT: Stuart, Susan G.  
; TITLE OF INVENTION: A NOVEL HUMAN JAK2 KINASE  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94304

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/196,480  
; FILING DATE:

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/567,508  
; FILING DATE: 05-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PP-0049US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-845-0555  
; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4482 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; LIBRARY: placenta  
; CLONE: 179527  
US-09-196-480-1

Alignment Scores:  
Pred. No.: 1.03e-233 Length: 4482  
Score: 2621.00 Matches: 532  
Percent Similarity: 66.03% Conservative: 193  
Best Local Similarity: 48.45% Mismatches: 342  
Query Match: 44.73% Indels: 31  
DB: 3 Gaps: 14

US-09-397-967A-16 (1-1099) x US-09-196-480-1 (1-4482)  
QY 20 SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl 39  
Db 447 TCCAGTTCCTCAGGT-----GTATCTTTACCATTCCTTTGGAAATCTGAGGACA 497  
QY 39 nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAla 59  
Db 498 TTATCTGACCTTTCCATCTGGGAGTATGTTGGAGAAGAAATCTGTATTGCTCTCTAA 557  
QY 59 sLaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe 79  
Db 558 AGCTTGTGTATCACACCTGTGTATCATATAATATGTTTCTTTAATGAGTGAACAGAAAG 617  
QY 79 rCysTrpPheProProSerHisIlePheCysIleGluAspValAlaAspThrGlnVal 99  
Db 618 GATCTGGTATCCACCACCAACCATGCTCTTCATATAGATGAGTCAACAGGACATAATGTACT 677  
QY 99 lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHis 117  
Db 678 CTACAGATATAGATTTTACTTTCCTCGTTGGTATTGTCAGTGGCAGCAACAGAGCCTACG 737  
QY 117 gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHis 137  
Db 738 GCATGGAATATCTCGAGGTGCTGAAGCTCCTCTTCTTGATGACTTTGTCTATGCTTACT 797  
QY 137 uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMet 157  
Db 798 CTTTGTCTCAGTGGCGCATGATTTTGTGATGGATGGATAAAAGTACCTGTGACTCATGA 857  
QY 157 sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGlu 177  
Db 858 AACACAGGAAGAATGCTTTGGGATGACAGTGTATGATATGATGATAGCAAGTAAAGAAA 917  
QY 177 nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProSe 197  
Db 918 CGATCAAAACCCACTGGCCATCTATACTCTATCAGCTACAGACATCTTACCAATG 977  
QY 197 rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgIleArgThrVa 217  
Db 978 TATTGCGACAAAGATCCAAGACTATCATATTTTGCAGAGGAAGCAATTAAGTACAGATT 1037  
QY 217 lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAla 237  
Db 1038 TCGCAGATTATTTCAGCAATTCAGCAATGCAAAAGCCACTCCCAAGAACTTGAACCTAA 1097  
QY 237 sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal 257  
Db 1098 GTATCTTATAAATCTGGAACTCTGCAGTCTGCCTTCTACACAGAAATTTGAAGTAA 1157  
QY 257 yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273  
Db 1158 AGAACCTGGAAGTGGTCTCTCAGGTGAGGAGATTTTGCACCACTATATAATACTGGA 1217  
QY 273 pAsnGlyIleProTrpSerSer-----AsnAspGluLe 284  
Db 1218 CGTGGAAATTCAGTGGTCAAGAGGGAACATAAGAAAGTGCAGACACTGACAGCAACAG 1277  
QY 284 uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304  
Db 1278 TTTACAGTTATATGCGATTTTCCATAATATATTGATGTCAGTATTGAAGCAAGCAACCA 1337  
QY 304 gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIle 324  
Db 1338 AGAGGGT---TCAAATGAAAGCCGAGTTGTAATCTATCCATAAGCAAGATGTTAAATCT 1394  
QY 324 uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGly 344  
Db 1395 GGAAATGGAACCTTAGCTCATTAAGGGAAGCTTTGTCTTCGTCAATTAATGATGATA 1454  
QY 344 rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProArgLe 364  
Db 1455 TTATAGATTAACTGCAGATGCATCATTTACCTCTGTAAAGAAAGTAGCACCTCCAGCCGT 1514



```

Db      3666  TGTAAATCAACGCCCTCTCTTAGGAGTACTAGTCTTCGAGTGATCAAAATA 3717
RESULT 9
US-09-972-800A-15
; Sequence 15, Application US/09972800A
; Patent No. 6534277
; GENERAL INFORMATION:
; APPLICANT: Hancock, W.
; APPLICANT: Ozkavnak, B.
; TITLE OF INVENTION: ROLES OF JAK/STAT FAMILY MEMBERS IN TOLERANCE
; TITLE OF INVENTION: INDUCTION
; FILE REFERENCE: 7853-192
; CURRENT APPLICATION NUMBER: US/09/972,800A
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US/09/549,654
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 5117
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (478)...(3876)
US-09-972-800A-15

Alignment Scores:
Pred. No.:      1,27e-233      Length:      5117
Score:          2621.00      Matches:     533
Percent Similarity: 66.12%      Conservative: 193
Best Local Similarity: 48.54%      Mismatches:  341
Query Match:      44.73%      Indels:       31
DB:                4          Gaps:          14

US-09-397-967a-16 (1-1099) x US-09-972-800A-15 (1-5117)

Qy      20  SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProGly 39
Db      585  TCCAGTCTTCAGGT-----GTATCTTTACCATTTCCCTTGGAAATCTGAGGCAGA 635
Qy      39  nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLy 59
Db      636  TTAATCTGACCTTCCATCTGGGGAGTATGTTGCAGAGAAATCTGATGCTGCTTCAA 695
Qy      59  sAlaCysGlyLeuLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSe 79
Db      696  AGCTTGTGGTATCACACCTGTGTATCATATAATATGTTTGTCTTAATGAGTGAACAGAAAG 755
Qy      79  rCysTrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa 99
Db      756  GATCTGGTATCCACCCACCATGCTTCCATATAGATGAGTCAACCGGCATATGTAAT 815
Qy      99  lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr 117
Db      816  CTACAGATAAGATTTTACTTCTCGTGTGATTCAGTGGCAGCAACAGGCCTATCG 875
Qy      117  gPheGlyLeuArgIlysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe 137
Db      876  GCATGGAATATCTCGAGGTGCTGAAGCTCTCTTCTTGTGATGACTTTTGTCTTACCT 935
Qy      137  uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy 157
Db      936  CTTTGTCTAGTGGCGGATGATTTTGTGTCAGGATGGATATAAAAGTACTGTGATCATGA 995
Qy      157  sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl 177
Db      996  AACACAGGAAGAAATGCTCTGGAGTGGCAGTGTAGATATGATGAGATAGCCAAAGAAA 1055
Qy      177  nAlaGlnArgProGlyGluLeuLeuHisThrValSerTyrIlysAlaCysLeuProProSe 197
Db      1056  CGATCAAAACCCCACTGGCCATCTATACTCTATCAGCTACAGACATTCCTTACCAAAATG 1115

```

```

Qy      197  rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrVa 217
Db      1116  TATTCGAGCAAAAGATCCAGACTATCATATATTTGACAGGAAGCGAATAAGGTACAGATT 1175
Qy      217  lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy 237
Db      1176  TCGCAGATTTATTTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAGCA 1235
Qy      237  sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl 257
Db      1236  GTATCTTATAAATCTGAAACTCTGCAGTCTGCCTTCTACACAGAGAAATTTGAAGTAA 1295
Qy      257  yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs 273
Db      1296  AGAACCTGGAAGTGGTCTCTCAGGTGAGGAGATTTTTCACCACTATAATAACTGGAAA 1355
Qy      273  pAsnGlyIleProTrpSerSer-----AsnAspGluLe 284
Db      1356  CGGTGGAATTCAGTGTCTCAGAGGGAACATAAAGAAAGTGAGACACTCACAGAACAGGA 1415
Qy      284  uPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProAr 304
Db      1416  TTTACAGTTATATTCGATTTTCTAATATATTATTCATGTCAGTATTACCAAGCAACCA 1475
Qy      304  gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe 324
Db      1476  AGAGGT---TCAATGAAAGCCGAGTGTGAATCTTCACTCAATAGCAAGATGGTAAAAATCT 1532
Qy      324  uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy 344
Db      1533  GGAAATTCAGTCTTATTAAGGAAGCTTTGCTTTCGTTGTCATTAAATGATGGGATA 1592
Qy      344  rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe 364
Db      1593  TTATAGATTAACTGCAGATGCATCATTCATCTCTGTAAGAGAGTAGACCTCCAGCCGT 1652
Qy      364  uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy 384
Db      1653  GCTTGAATAATACAAAGCAACTGTCATGGCCCAATTTTCGATGATTTTGCATTAGTAA 1712
Qy      384  sLeuLyAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTy 404
Db      1713  ACTGAGAAAGACAGTAAATCAGACTGGACTGTATGTAATCTTCGATGCACTCTTAGGACTT 1772
Qy      404  rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCy 424
Db      1773  TAATAAATATTTTGTGACTTTTGTGCTGCGAGCGAGAAAATGCTCATTAATATAACACTG 1832
Qy      424  sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr 444
Db      1833  TTTGATTACAAAATAAGATGAAGATACACCTCAGTGGGACAAAGAAAGAACTTCAG 1892
Qy      444  gSerLeuArgGluLeuLeuAlaAlaCysTyrAsnSerGlyLeuArgValAspGlyAlaAl 464
Db      1893  CAGTCTTTAAGATCTTTTGAATTTGTTACAGATGGAAACTGTTCTCGTCAGACAAATATA 1952
Qy      464  aLeuTyrLeuThrSerCysAlaProArgProLysGluLysSerAsnLeuIleValVa 484
Db      1953  TTTCCAGTTTACTAAATGCTGTCCCCCAAGGCCAAAGAGATAAATCAAACTTCTAGTCTT 2012
Qy      484  lArg-----ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysAlaLe 503
Db      2013  CAGAACGAATGGTGTCTCTGATGTACCAACTCACCAACATTAACAGAGGCTCTACTCAT 2072
Qy      503  uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGl 523
Db      2073  GAACCAAAATGGTGTTCACAAAATCAGAAATGAAGATTTTGATATTTAATGAAGCCTTGG 2132
Qy      523  yHisGlySerPheThrIlysIlePheArgGlySerArgArgGluValValAsp---GlyGl 542
Db      2133  CCAAGGCACCTTTTCAAAAGATTTTAAAGCGCTACGAAGAGAGATAGGAGACTACGGTCA 2192

```







; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 94..3480  
 US-08-665-574C-8

## Alignment Scores:

Pred. No.: 1,01e-233 Length: 3629  
 Score: 2619.50 Matches: 537  
 Percent Similarity: 65.50% Conservative: 190  
 Best Local Similarity: 48.38% Mismatches: 350  
 Query Match: 44.70% Indels: 33  
 DB: 3 Gaps: 15

US-09-397-967a-16 (1-1099) x US-08-665-574C-8 (1-3629)

```

Qy 4 ProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSerGluAla 23
Db 145 CCTGTACATCAGATGCTGATATCTCTGGAAGTGTCTAATCTGTGAAGCAGATAGGCCA 204
Qy 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43
Db 205 GTCCTTCAAGTGTATCTGTACCATCTCTTGGCAAGCTGAAGGAGATATCTGAAGTTT 264
Qy 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
Db 265 CCAAGTCGAGATGTTGTCAGAGAATTTGTGTGGCTGCTTCTAAGCTTTGGTATT 324
Qy 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
Db 325 ACGCCTGTGTATCATATATGTTTGGTGAATGATGAGTGAACCGAAGGATCTGTATCCCA 384
Qy 84 ProSerHisLeuPheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103
Db 385 CCCAATCATGTCTTCCATAGACAGAGTCAACACGAGCATGACATCTACAGGATAAGG 444
Qy 104 PheTyrPheProAspTyrPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121
Db 445 TTCTACTTCCCTCATTTGTAAGTGTAGTGGAGCAGCAGACCTACAGATACGAGTGTCC 504
Qy 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
Db 505 CGTGGGCTCAAGTCTCTGCTGTATGATCTTGTCTATGCTTACCTTTTGTCTACGTGG 564
Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161
Db 565 CGGCATGATTTGTTCACGATGATGATAAAGTACCTGTGACTCATGAAACTCAGGAAGAG 624
Qy 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181
Db 625 TGCTCTGGGATGCGGTGTAGATCATGATGAGATAGCTAAGGAGAAGACCCAGACTCCA 684
Qy 182 GlyGluLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
Db 685 CTGGCTGTCTATAACTCTGTCTGAGTCAAGACATCTTACCAAGTCGTCGCGCGAAG 744
Qy 202 IleGlnGlyGlnAsnPheValThrArgArgIleArgThrValValLeuAlaLeu 221
Db 745 ATCCAAAGACTATACATATTTAAACCGGAGCGAATCAGGTACAGATTTTCGAGATTCATT 804
Qy 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241
Db 805 CAGCAATTCAGTCAATGTAAGCCACTGCCAGGAACCTAAACCTTAAGTATCTTATAAAC 864
Qy 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal----- 256
Db 865 CTGGAAACCTCGCATCTGCTCTCTACACAGAACAGTGTGAAGTAAAGAAATCTGCAAGA 924
Qy 257 GlyLeuProGlyAlaGlnGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIle 276
Db 925 GGT---CCTTCAGGTGAGGAGATTTTTCGAACCATATATAAATTAATGGAACCGTGAATT 981

```

```

Qy 277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287
Db 982 CAGTGGTCAAGAGGGAACAATAAGAAAGTGAGACACTGACAGACAGCAGCTACAGTTA 1041
Qy 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307
Db 1042 TATTGTGATTTCCCTGATATATTGATGTCTAGTATTAAAGCAAGCAACCAG---GAATGC 1098
Qy 308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327
Db 1099 TCAAAATGAAGTAGAATTGTAATCTGCATAAACAAGATGTAAGTTTTCGAGATAGAA 1158
Qy 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
Db 1159 CTTAGCTCAITTAAGAGAGCCCTTGTCTTGTCTATTAATTGACGGGTATTACAGACTA 1218
Qy 348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuGluGlu 367
Db 1219 ACTGGGATGCGCACCATTACCTCTGCAAGAGAGTGGCTCCCCAGCTGTGCTCGAGAAC 1278
Qy 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla 387
Db 1279 ATACACAGCAACTGCCACGCGCCCAATATCAATGGATTTTGCCATTAGCAACTAAGAAG 1338
Qy 388 AlaGlySerLeuProGlyThrTyrIleLeuArgSerProGlnAspTyrAspSerPhe 407
Db 1339 GCGGGTAAACAGACTGAGACTATATGTCTACGATGCAGCCCTAAGGACTTCAACAAATAC 1398
Qy 408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
Db 1399 TTCTGACCTTTGCTGTGTGAGCGAGAAATGTCATTGAATATAAACAACCTGTTGATTACG 1458
Qy 428 GluAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
Db 1459 AAGAATGAGAATGGAGAATCAACCTCAGCGGACTAAGAGGAACCTTCAGTAACCTTAAG 1518
Qy 448 GluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467
Db 1519 GACCTTTTGAATTTGCTACCATGAGTGAACCTGTGCGCTCAGACAGATATCATCTTCCAGTTT 1578
Qy 468 ThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgArgGly 487
Db 1579 ACCAATGTGTCCTCCCAAGCCAAAGATAAATCAAACTTCTCGTCTTCAGA----- 1632
Qy 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504
Db 1633 ACAATATGTTATTTCTGATGTTCTAGATCTCACCACATTTACAGAGGCATAATAATGTGAAT 1692
Qy 505 GluLeuSerPheHisThrIleProThrAspSerLeuGluThrHisGluAsnLeuGlyHis 524
Db 1693 CAAATGTTGTTTCAAAAATCAGAAATGAAGATTTTAAATTTAATGAAGTCTTGGCCAA 1752
Qy 525 GlySerPheThrIleLysIlePheArgGlySerArgGluValValAsp---GlyGluThr 543
Db 1753 GGTACTTTTACAAAATTTTAAAGGTGTAAAGAGAGAGTTGGAGATTATGTTCAACTG 1812
Qy 544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
Db 1813 CACAAAACGGAAGTCTTTTGAAGTCTTAGATAAAGCACATAGGAACATTATCAGAGTCT 1872
Qy 564 PheLeuGluAlaLysSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583
Db 1873 TTCTTCAAGACGACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932
Qy 584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602
Db 1933 GGTGTCGTGTCTGTGAGAGGAGAACTTCTGTTCAAGAAATTTGTAATAATTTGATCA 1992
Qy 603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnVal 622
Db 1993 CTGGATACATACCTGAAGAGAACAAATTCATAAATATATATATGGAACCTTGAGGTG 2052

```



Qy 623 ThrLysGlnLeuAlaValAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsn 642  
Db 2053 GCTAAGCAGTGGCCATGGCCATGATTTCTAGAGAAATATCCCTATTTCATGGGAT 2112  
Qy 643 ValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnProPro 660  
Db 2113 GTGTGTCTAAATAATATCTCTTATCAGAGAAGAAGACAGAGAACCGGGAACCCACCT 2172  
Qy 661 PheLeuLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680  
Db 2173 TTTCAACAACTTAGTGATCTCTGGCCTCTAGCATTTACATTTCTACCGAAGACATTTCTCAG 2232  
Qy 681 AspArgLysProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeuSerLysLeuGlu 700  
Db 2233 GAGAGAAATACCAAGGTACCTCTGAAATGATGAGAAATCTTAAATAATCTCAATCTGGCA 2292  
Qy 701 AlaAspLysTyrGlyPheGlyAlaThrThrTyrGluValPheGlnArgGlyProLahis 720  
Db 2293 ACAGACAAAGTGGAGCTTGGGACCACTCTGTGGAGATCTGCAGTGGAGAGATAAGCCC 2352  
Qy 721 IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740  
Db 2353 CTGAGTGTCTGATCTCTCAAGAAAGCTGCGATCTCTATGAGATGAAGCATCAGCTTCTCT 2412  
Qy 741 AlaLeuLysTyrThrGluLeuAlaGlyLeuLysLeuThrGlnCysMetAlaTyrAspProGly 760  
Db 2413 GCACCAAGTGGACAGATTGCAAACTTATAATAATTTGATGGACTATGAGCCAGAT 2472  
Qy 761 ArgArgProSerPheArgAlaLeuLeuArgAspLeuAsnGlyLeuLysLeuThrSerAspTyr 780  
Db 2473 TTTGAGGCTGCTTTCAGAGCTGTCATCGTCTCTTAAACAGCTGTTTACTCCAGATTAT 2532  
Qy 781 GluLeuLeuSerAspProThrProGlyLysLeuProSerProArg---AspGluLeuCysVal 799  
Db 2533 GAACCTACTAACGAA---AATGACATGCTACCAACATCAGATAGTGGCCCTAGGGTTT 2589  
Qy 800 AlaGlyValAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluAlaGlnLys 819  
Db 2590 TCTGGTGCT-----TTGAAAGACAGGACCTTACACAGTTTGAAGAGAGACACTTGAAG 2643  
Qy 820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuLysCysArgTyrAspPro 839  
Db 2644 TTTTCTACAGCAGCTTGGCAAGGTAATCTCGGAGTGTGGAGATGTGCGGCTATGACCCG 2703  
Qy 840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859  
Db 2704 CTGACGACAACTGCGGAGGTGGTGGTGTGAAGAACTCCAGACAGACACTGGAAG 2763  
Qy 860 GlnGlnArgAspPheGlnArgGluLysGlnIleLeuLysAlaLeuHisSerAspPheIle 879  
Db 2764 CACCTCGGACCTTTGAGAGGAGATCGAGATCTGAAATCTTGCAGCATGACACATC 2823  
Qy 880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899  
Db 2824 GTCAAGTACAAGGAGGTGCTCAGTGCCTGCGGCGGCGCAACCTAAGATTAAATATGGA 2883  
Qy 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr 918  
Db 2884 TATTTACCATATGGAAGTTTACGAGACTATCTCCAAATAAATAAGAACCGATAGATCAC 2943  
Qy 919 AspArgLeuLeuLeuPheAlaTyrGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 938  
Db 2944 AAAAACTTCTTCAATACATCTCAGATATGCAAGGCGATGGAATATCTTGTACAAA 3003  
Qy 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGlnSerGluAlaHisVal 958  
Db 3004 AGGTATATCCAGGAGCTGCGCAACAGGAACATATTTGTGGAATAATGAGACAGGGTT 3063  
Qy 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978  
Db 3064 AAAATAGGACTCTCGGATTAACCAAGCTTTGCGGAGGACAAAGAAATACTACAAGTA 3123  
Qy 979 ArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAsnIle 998

Db 3124 AAGGAGCCAGGGAAGCCCATATTTCTGTACGACCTGAATCTTACGAGAGCAAG 3183  
Qy 999 PheSerArgGlnSerAspValTyrSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
Db 3184 TTTTCTGTGGCTCAGATGTGTGGAGCTTTGGAGTGGTCTATATGCAACTTTTCATAC 3243  
Qy 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038  
Db 3244 ATCGAGAGAGTAAAGTCCACCCGTCGGAATTTATGCAATGATTGGCAATGATAACAA 3303  
Qy 1039 GlyProProLeuCys---ArgLeuLeuGlnLeuLeuAlaGluGlyArgArgLeuProPro 1057  
Db 3304 GGGCAATGATGTGTCTCCATTTGATAGAGCTACTGAAGCAACCGAAGATTGCCAAG 3363  
Qy 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTyrAlaProGluPro 1077  
Db 3364 CAGAAGAGTCCAGATGAGATTATGTGATCATGACAGAGTGTGGACACACATGTG 3423  
Qy 1078 HisAspArgProAlaPheAlaThrLeuSer 1087  
Db 3424 ACCGAGCTCCCTCCTTCAGGACCTTCG 3453  
RESULT 12  
US-08-946-994-8  
; Sequence 8, Application US/08946994  
; Patent No. 6210654  
; GENERAL INFORMATION:  
; APPLICANT: Ihle, James N.  
; APPLICANT: Silvennoinen, Ollie  
; APPLICANT: Mitthum, Bruce A.  
; APPLICANT: Quelle, Frederick W.  
; TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal  
; TITLE OF INVENTION: Transduction  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1100 New York Avenue, Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3934  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/946,994  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/665,574  
; FILING DATE: 18-JUN-1996  
; APPLICATION NUMBER: 08/282,012  
; FILING DATE: 29-JUL-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/097,997  
; FILING DATE: 29-JUL-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/118,968  
; FILING DATE: 09-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fox, Samuel L.  
; REGISTRATION NUMBER: 30,353  
; REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3629 base pairs  
; TYPE: nucleic acid



QY	277	ProTrpSerSer-----AsnAspGluLeuPheGlnThr	287
DB	982	CAGTGGTCAAGAGGGAACATAAGGAAGTGCAGACATGACAGAACAGGACGTCACACTTA	1041
QY	288	PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro	307
DB	1042	TATTGTGATTTCCCTCGATATTATTGATGTGCAGTATTAAAGCAAGCAAAACCCAG--GAATGC	1098
QY	308	AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu	327
DB	1099	TCAAAATGAAGTAGAATTTGTAACGTCCATAAACAGATGGTAAGTTTTTGGAGATGAGA	1158
QY	328	PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu	347
DB	1159	CTTAGCTCAATTAAGAAGACCTTGTCTATTTCGTGTCATTAAATTCACGGTATTACAGACTA	1218
QY	348	IleCysAspSerArgHisTyrPheCysLeuGluValAlaProProArgLeuLeuGluGlu	367
DB	1219	ACTCGGATCGCACCAATTACCTCTGCAAGAGAGGTGCTCCCCAGCTGTGTCGAGAAC	1278
QY	368	GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla	387
DB	1279	ATACACAGCAACTGCCACGCCCAATATCAATGGATTTTGCATTAGCAAACTAAAGAAG	1338
QY	388	AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe	407
DB	1339	CGCGGTAACCAAGACTGGACTATATGTCTACGATCGAGCCCTAAGGACTTCAACAATAC	1398
QY	408	LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg	427
DB	1399	TTTCTGACCTTTGCTGTGAGCGAGAAAATGTCATTGAATATTAACACTGTTTGATTACG	1458
QY	428	GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg	447
DB	1459	AAGAATGAGAAATGGAGAAATACACCTCAGCGGACTTAAGAGGAACCTTCAGTACACCTTAAG	1518
QY	448	GluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu	467
DB	1519	GACCTTTTGAATTTGCTACCAAGATGGAAACTGTGCGCTCAGACAGTATCATCTTCCAGTTT	1578
QY	468	ThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgArgGly	487
DB	1579	ACCAAAATGCTGCCCCCAAGCCAAAAGATAATCAAACCTTCTCGTCTCTCAGA-----	1632
QY	488	CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr	504
DB	1633	ACAAATGGTATTCTCATGTTTCAGATCTCACCACCAATTACAGAGGCAATAATATGTGAAT	1692
QY	505	GlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGlyHis	524
DB	1693	CAAAATGGTGTTCACAAAATCAGGAATGAAGATTTAATATTAAATGAAGTCTTGGCCAA	1752
QY	525	GlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp---GlyGluThr	543
DB	1753	GGTACTTTTACAAAAATTTTAAAGGTGTAAAGAGAAAGTTGGAGATTTATGGTCAACTG	1812
QY	544	HisAspSerGluValLeuLysValMetAspSerArgHisArgAsnCysMetGluSer	563
DB	1813	CACAAAACGGAGTCTTTTGAAGATCCCTAGATAAAGCACATAGGAACATTTCAGAGTCT	1872
QY	564	PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis	583
DB	1873	TTCTTCGAAGCAGCAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1932
QY	584	GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla	602
DB	1933	GGTGTCTGTGCTGTGGAGGAGGAACATCTCGTTCAAGAATTTGTAAAAATTTGGATCA	1992
QY	603	IleAspMetTyrLeuArgLysArgGlyHisIleuValSerAlaSerTrpLysLeuGlnVal	622
DB	1993	CTGGATACATACCTGAAAGAGAACAAAATTCATAAAATATATTATGGAACATTTGGAGTG	2052
QY	623	ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAsnIleValGlyLeuProHisIleVal	642

Db	2053	GCTAAGCAGTGGCAGGCGCCATGATTTCTAGAGAAATCCCTATTATCGGAT	2112	Db	3124	AAGGAGCCAGGGAAGCCCATATTTCTGGTACGACCTGAAATCCTTGAGGAGCAAG	3183
Qy	643	ValSerAlaArgLysValLeuLeuAlaArgGluGlyGly- - - - - AspGlyAsnProPro	660	Qy	999	PheSerArgGlnSerAspValTrpSerPheGlyValValLeuLeuPheThrTyr	1018
Db	2113	GTGTGCTGCTAAATATCTCTGCTTATCAGAGAAGAAGCAGGAGAACGGGGAACCCACCT	2172	Db	3184	TTTTCTGTGGCTCAGATGTGTGGAGCTTTGGAGTGGTCTATACGAACTTTTCATAC	3243
Qy	661	PheLeuLysLeuSerAspProGlyValSerProThrValLeuSerLeuMetLeuThr	680	Qy	1019	CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu	1038
Db	2173	TTTCATCAAACTTAGTGATCTCTGGCATTTAGCATTTACAGTTCTACCGAAGGACATTTCTAG	2232	Db	3244	ATCGAGAAGAGTAAAGTCCACCCCGTGGAAATTTATGCGAATGATTGGCAATGATAACAA	3303
Qy	681	AspArgLysProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu	700	Qy	1039	GlyProProLeuCys- - - ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro	1057
Db	2233	GAGAAATACCACTGGTCTCTGTAATGATGAGATCTTAAATCTCAATCTGGCA	2292	Db	3304	GGGCAATGATGTGTTCATTTGATAGAGCTACTGAGAGCAACGGAAGATTGCCAAG	3363
Qy	701	AlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAlaHis	720	Qy	1058	ProProThrCysProThrGluValGlnLeuMetGlnLeuCysTrpAlaProGluPro	1077
Db	2293	ACAGACAAGTGGAGCTTCGGGACCACTCTGTGGAGATCTGCAGTGGAGGAGATAAGCC	2352	Db	3364	CCAGAGGATGCCAGATGAGATTTATGTGATCATGACAGAGTGTCTGGAACCAATGTG	3423
Qy	721	IleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro	740	Qy	1078	HisAspArgProAlaPheAlaThrLeuSer	1087
Db	2353	CTGAGTCTCTGATTTCTCAAAAGAGCTGCAGTTCTTGAAGATAAGCATCAGCTTCT	2412	Db	3424	AGCCAGCGTCCCTCTTCAGGACCTTTGCG	3453
Qy	741	AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly	760				
Db	2413	GCACCAAGTGGACAGATTAGCAACCTTATAAATTAATGATGAGTATGAGCCAGAT	2472				
Qy	761	ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr	780				
Db	2473	TTTCAGGCGCTTTCAGAGCTGTCATCCGCTGATCTTTAAACAGCCTGTTTATCTCCAGAT	2532				
Qy	781	GluLeuLeuSerAspProThrProGlyIleProSerProArg- - - AspGluLeuCysVal	799				
Db	2533	GAACCTACTAACGAA- - - AATGACATGCTACCAAAACATGAGATAGTGCGCCCTAGGGTT	2589				
Qy	800	AlaGlyValGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys	819				
Db	2590	TCGTGGTCT- - - - - TTGGAAGCAGGACCCCTACACAGTTTGAAGAGACACTTGAAG	2643				
Qy	820	TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro	839				
Db	2644	TTTCTACAGCAGCTTGGCAAGGTAACTTCGGAGTGTGGAGATGTGCCGCTATGACCCG	2703				
Qy	840	LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp	859				
Db	2704	CTGCAGGACAACTGCGGAGTGTGCTGTGAAGAACTCCAGCACAGCAGCTGAAGAG	2763				
Qy	860	GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle	879				
Db	2764	CACCTCCGAGCTTTGAGGGAGATCGAGATCTGAAATCCTTGCAGCATGACAAATC	2823				
Qy	880	ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu	899				
Db	2824	GTCAAGTACAAGGAGTGTGCTACAGTGGGTCCGGGCAACCTTAAGATTAAATTTATGAA	2883				
Qy	900	TyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArg- - - GlyLeuHisThr	918				
Db	2884	TATTTACCATATGGAAGTTTACGAGACTATCTCAAAACATAAAGAACGGATAGATCAC	2943				
Qy	919	AspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg	938				
Db	2944	AAAAAATTTCTCAATACATCTCAGATATGCAAGGCGCATGGAATATCTTGGTACAAA	3003				
Qy	939	ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal	958				
Db	3004	AGGTATATCCAGGAGCTGGCAACAGGAACATTTGTGTGAAATGAGAACAGGGTT	3063				
Qy	959	LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal	978				
Db	3064	AAAAATAGGAGACTTCGGATTAAACCAAGTCTTGGCGAGGCAAGAAATATCAAAAGTA	3123				
Qy	979	ArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAsnIle	998				

Alignment Scores: 2.9e-231 Length: 4078  
 Pred. No.: 2594.00 Matches: 531  
 Score:

RESULT 13  
 US-09-016-434-297  
 ; Sequence 297, Application US/09016434  
 ; Patent No. 6500938  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Janice Au-Young  
 ; APPLICANT: Jeffrey J. Seilhamer  
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 ; NUMBER OF SEQUENCES: 1490  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 ; STREET: 3174 PORTER DRIVE  
 ; CITY: PALO ALTO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/016,434  
 ; FILING DATE: HEREWITH  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Zeller, Karen J.  
 ; REGISTRATION NUMBER: 37,071  
 ; REFERENCE/DOCKET NUMBER: PA-0002 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (650) 855-0555  
 ; TELEFAX: (650) 845-4166  
 ; INFORMATION FOR SEQ ID NO: 297:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 4078 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: PLACNOB01  
 ; CLONE: 179527  
 ; US-09-016-434-297

Percent Similarity:	65.94%	Conservative:	193
Best Local Similarity:	48.36%	Mismatches:	343
Query Match:	44.27%	Indels:	32
DB:	4	Gaps:	14

US-09-397-967A-16 (1-1099) x US-09-016-434-297 (1-4078)

Qy	20	SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGl	39
Dd	447	TCCAGTTCTTCCTCAGT-----GTATCTTTACCATTCCTTGCGAAATCTGGACGACA	497
Qy	39	nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaI	59
Dd	498	TTATCTGACCTTTCCAICTGGGGAGTAGTTCGAGAAGAAATCTGTATTGCTGCTCTCAA	557
Qy	59	sAlaCysGlyLleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheS	79
Dd	558	AGCTTGTTGGTATCACACTGTGTATCATATAATATGTTCTTTAATGAGTGAACAGAAA	617
Qy	79	rCysTrpPheProSerHisLlePheCysLleGluAspValAspThrGlnValLeuVa	99
Dd	618	GATCTGGTATCCACCACCAATGCTTCATATAGATGAGTCAACGACGCATAATGTACT	677
Qy	99	LtyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisAr	117
Dd	678	CTACAGAAATAAGATTTTACTTTCCTCG-TGGTATTGCCAGTGGCAGCAACAGAGCCTATCG	736
Qy	117	gPheGlyLeuArgLysAspLeuThrSerAlaLleLeuAspLeuHisValLeuGluHisL	137
Dd	737	GCATGGAAATATCTCGAGGTGCTGAAGCTCTCTTCTTGATGACTTTGTGATGCTTACTCT	796
Qy	137	uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLy	157
Dd	797	CTTTGTCTCAGTGGCGGATGATTTGTGTCATGGATGGATAAAGTACCTGTGACTCATGA	856
Qy	157	sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetalArgGluGI	177
Dd	857	AACACAGGAAGAATGCTTCGGATGACAGGTGTAGATATGATGAGAATAGCAAGAAAA	916
Qy	177	nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe	197
Dd	917	CGATCAAAACCCACTGGCCATCTATACTCTATCAGCTACAGACATCTTTACCAACATG	976
Qy	197	rLeuArgAspValLleGlnGlyGlnAsnPheValThrArgArgLleArgArgThrVa	217
Dd	977	TATTCGACGAAAGATCCAAAGACTATCATATTTTGACAGGAAGCGAATAAGGTACAGATT	1036
Qy	217	lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy	237
Dd	1037	TCCGAGATTATTTCAGCAATTCAGCCAAATGCAAGCCACTGCCAGAACCTTTGAACCTTAA	1096
Qy	237	sTyrlLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGI	257
Dd	1097	GTATCTTATAAATCTGGAAACTCTGCAAGTCTGCCCTTCTACACAGAGAAATTTGAAGTAA	1156
Qy	257	yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs	273
Dd	1157	AGAACCTGGAAAGTGGTCTTCAGGTGAGGAGATTTTTTGCAACCATTAATACTGAGAAA	1216
Qy	273	pAsnGlylleProTrpSerSer-----AsnAspGluLe	284
Dd	1217	CGGTGGAATTCAGTGGTCAAGAGGGAAACATAAAGAAAGTCAGACACTGACAGAACAGGA	1276
Qy	284	uPheGlnThrPheCysAspPheProGluLleValAspValSerlleAsnGlnAlaProAr	304
Dd	1277	TTTACAGTTATATGGGATTTTCTAATATATTGATGTGATGATTAAGCAAGCAAAACCA	1336
Qy	304	gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisLleLe	324
Dd	1337	AGAGGT---TCAATGAAGCCGAGTTGTAACATATCATTAAGCAAGATGGTAAAAATCT	1393
Qy	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy	344

1394	GGAAATTTGAACCTTAGCTCATTAAAGGGAAGCTTTGTTCTTTCGTGTCATTAAATTTGATGGATA	1453
344	rPheArgLeuIleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLe	364
1454	TTATAGATTTAACTGCAGATGCACATCACTCTGTAAAGAAGTAGCACCCTCCAGCCGT	1513
364	uLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLy	384
1514	GCTTGAAATATACAAAGCAACTGTCATGGGCCCAATTCGATGATTTTGCCATTAGTAA	1573
384	sLeuIysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTy	404
1574	ACTGAAGAAACAGAGTAATCAGACTGCTATGTACTTCGTATGCAGTCAGTCTCTAAGGACTT	1633
404	rAspSerPheLeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrIlysGlyCy	424
1634	TAATAAATATTTTGACTTTTGCTGTGCGAGGAGAAATGTCAATGAATATAACAACCTG	1693
424	sLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisAr	444
1694	TTTGATTACAAAAAATGAGATGAAGAGTAGTACAACCTCAGTGGGACAAAGAAGAACTTCAG	1753
444	sSerLeuArgGlnLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGlyAlaAl	464
1754	CAGTCTTTAAAGATCTTTTGAATTGTTACCAGATGGAAACTGTTCGTGCAGACAAATAAT	1813
464	aLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValva	484
1814	TTTCCAGTTTACTAAATGCTGTCCCCCAAGGCCAAAGATATAATCAACCTCTTAGTCTT	1873
484	lArg---ArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAlaLe	503
1874	CAGAACGAATGGTCTTCTCATGTATACCAACCTCACCAACATTACAGAGGCTACTCATAT	1933
503	uThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeuGl	523
1934	GAACCAATGGTGTTCACAAAATCAGAAATGAAGATTTTGATATTTAATGAAGACCTTGG	1993
523	yHisGlySerPheThrIysIlePheArgGlySerArgArgGluValValAsp---GlyGl	542
1994	CCAAGGCACITTTACAAGATTTTAAAGCGTACGAAGAGAAGTAGGAGACTTACCGTCA	2053
542	uThrHisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGl	562
2054	ACTGCATGAACACAGAAGTCTTTTAAAGTTCTGGATAAAGCACACAGGAACATTATTCAGA	2113
562	uSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLe	582
2114	GTCTTTCTTTGAAGCAGCAAGTAGTAGTACCAAGCTTTCTCACAAGCATTTGGTTTAAA	2173
582	uHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyrLeuGl	601
2174	TTATGGAGTATGTCTGTGGAGCAGCAGATATTCCTGGTTTCAGAGTTTGTAAATTTGG	2233
601	yAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuGl	621
2234	ATCACTAGATACATATCTGAAAAAGAAATAAAATTTGTATAAATATATTATGGAACCTTGA	2293
621	nValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGl	641
2294	AGTTGCTAAACAGTTGGCATGGGCCATGCAATTTCTTAGAAGAAAAACACCCCTTATTATGG	2353
641	yAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnPr	659
2354	GAATGATATGGCCCAAAATATCTCGCTTATCAGAGAAGAAAGACAGGAAGACAGGAAATCC	2413
659	oProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetIle	679
2414	TCCTTTTCATCAAACTTAGTGATCTCTGGCATTAGTATTACAGTTTTCGCCAAGGACATCT	2473
679	uThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLe	699
2474	TCAGAGAGAATAACCATGGGTACCACTGAATGCATTTGAAAATCTCAAAAATTTAAATTT	2533

QY	699	uCluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAl	719	1056	oProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl	1076
Db	2534	GCACACACAAATGGAGTTTGGTACCACTTTGGGGAATCTCGAGTGGAGAGATAA	2593	3605	AAAGACAGATGATCCCGAGATCTATATGATCATGACAGATGCTGAACAATAA	3664
QY	719	aHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnIle	739	1076	uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu	1093
Db	2594	ACCTCTAAGTCTCTGGATTCTCAAGAAAGCTACAAATTTTATGAAGATAGGCATCAGCT	2653	3665	TGTAATAACACGCCCTCTCTTAGGATCTAGCTCTTGAGTGGATCAATAA	3716
QY	739	uProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPr	759			
Db	2654	TCCTGCACCAAGTGGCAGAAATTAGCAAAACCTTATAATAATTGTATGGATTATGAACC	2713			
QY	759	oGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAs	779			
Db	2714	AGATTTTCAGCCCTCTCTTCAGAGCCCATCATACAGATCTTAACAGTTTGTACTCCAGA	2773			
QY	779	pTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCy	798			
Db	2774	TTATGAACCTATTAAACAGAA---AATGACATGTTTACCAATATATGAGTATAGTGGC	2830			
QY	798	sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisIle	818			
Db	2831	GTITTTCTGTGTGCC-----TTTGAAGACCGGGATCCTCACACAGTTTTCAGAGAGACATTT	2884			
QY	818	uLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAs	838			
Db	2885	GAAATTTCTACACCACTTGGCAAGGTATATTTGGAGGTGTGNAGATGTCGCGTATGA	2944			
QY	838	pProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPr	858			
Db	2945	CCCTCTACAGGACACACATGGCGAGTGGTGGCTGTAAAAAGCTTCAGCATAGTACTGA	3004			
QY	858	oAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPh	878			
Db	3005	AGAGCACCTTAAGAGACTTTGAAAGGGAATTTGAAATCCCTGAAATCCCTACAGCATGACAA	3064			
QY	878	eIleValLysTyrArgLysValSerTyrGlyProGlyArgGlnSerLeuArgGluValMe	898			
Db	3065	CATTGTAAAGTACAGGAGGTGTCTACAGTGTGGTGGCGTAACTTAAATAATATAT	3124			
QY	898	tGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHi	917			
Db	3125	GGATATTTTACCATATGGAAGTTTACGAGACTATCTTCAAAAAACATAAAGACGGTAGA	3184			
QY	917	sThrAspArgLeuLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyrIleuGlyAl	937			
Db	3185	TCACATAAAACCTTCTGCAGTACACATCTCAGATATGCAAGGGTATGAGATATCTTGTAC	3244			
QY	937	aArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHi	957			
Db	3245	AAAAGGTATATCCACAGGATCTGGACAGAGAAATATATTTGGTGGAGACGAGACAG	3304			
QY	957	sValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrIyVa	977			
Db	3305	AGTTTAAATTTGGRGATTTGGTTTAAACCAAGTCTTCCACAGACAAAGAACTACTATAA	3364			
QY	977	lValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAs	997			
Db	3365	AGTAAAAAGAACCTGGTGAAGTCCCATATTTCTGGTATGCTCCAGAAATCATCTGACAGAG	3424			
QY	997	nIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheTh	1017			
Db	3425	CAAGTTTCTGTGGCCTCAGATGTTTGGAGCTTTGGAGTGGTCTGTATGAACCTTTTCAC	3484			
QY	1017	rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluAr	1037			
Db	3485	ATACATTGAGAAGAGTAAAGTCCACAGCGGAATTTATGCGTATGATGTCGAATGACAA	3544			
QY	1037	gGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgArgLeuPr	1056			
Db	3545	ACAAGGACAGATGATCGTGTCTTCCATTTTGATAGAACTTTTGGAAGAAATATGGAAGATTACC	3604			
QY				1056	oProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGl	1076
Db				3605	AAAGACAGATGATCCCGAGATCTATATGATCATGACAGATGCTGAACAATAA	3664
QY				1076	uProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu	1093
Db				3665	TGTAATAACACGCCCTCTCTTAGGATCTAGCTCTTGAGTGGATCAATAA	3716
RESULT 14						
US-08-446-038B-2						
Sequence 2, Application US/08446038B						
Patent No. 5658791						
GENERAL INFORMATION:						
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;						
APPLICANT: Harpur, Ailsa						
TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase						
NUMBER OF SEQUENCES: 23						
CORRESPONDENCE ADDRESS:						
ADDRESSEE: Felfe & Lynch						
STREET: 805 Third Avenue						
CITY: New York City						
STATE: New York						
COUNTRY: USA						
ZIP: 10022						
COMPUTER READABLE FORM:						
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage						
COMPUTER: IBM PS/2						
OPERATING SYSTEM: PC-DOS						
SOFTWARE: Wordperfect						
CURRENT APPLICATION DATA:						
APPLICATION NUMBER: US/08/446,038B						
FILING DATE: 19-MAY-1995						
CLASSIFICATION: 424						
PRIOR APPLICATION DATA:						
APPLICATION NUMBER: 08/064,067						
FILING DATE: 30-Jun-1993						
APPLICATION NUMBER: PCT/US91/08889						
FILING DATE: 26-No. 5658791-1991						
PRIOR APPLICATION DATA:						
APPLICATION NUMBER: Australian PK3594/90						
FILING DATE: 28-No. 5658791-1990						
PRIOR APPLICATION DATA:						
APPLICATION NUMBER: Australian 88229/91						
FILING DATE: 27-No. 5658791-1991						
ATTORNEY/AGENT INFORMATION:						
NAME: Hanson, No. 5658791man D.						
REGISTRATION NUMBER: 30,946						
REFERENCE/DOCKET NUMBER: LUD 5244						
TELECOMMUNICATION INFORMATION:						
TELEPHONE: 212-688-9200						
TELEFAX: 212-838-3884						
INFORMATION FOR SEQ ID NO: 2:						
SEQUENCE CHARACTERISTICS:						
LENGTH: 3495 base pairs						
TYPE: nucleic acid						
STRANDEDNESS: single						
TOPOLOGY: linear						
MOLECULE TYPE: nucleic acid						
US-08-446-038B-2						
Alignment Scores:						
Pred. No.: 1,44e-209						
Score: 2360.00						
Percent Similarity: 65.89%						
Best Local Similarity: 49.34%						
Query Match: 40.27%						
DB: 1						
Length: 3495						
Matches: 489						
Conservative: 164						
Mismatch: 300						
Indels: 38						
Gaps: 15						
US-09-397-967A-16 (1-1099) x US-08-446-038B-2 (1-3495)						
QY	128	IleLeuAspLeuHisValLeuGluHisPheAlaGlnHisArgSerAspLeuValSer	147			
Db	1	CTGCTTGTAGCTTTGTTCATGCTCTTACCTTCCCTCAGTGGCGCATGATTTTGTTCAC	60			

148 GlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGluPheLeuSerLeuAlaVal 167  
1129 GTTACAGCTCTCACCACCAATATTAGTGAATCAATGGTGTTCACCAA 1188  
511 IleProThrAspSerLeuGluTriPheHisGluLeuGlyHisGlySerPheThrLysIle 530  
1189 ATCAGGAATGAAGATTTCATATTATTAATGAAGCCTTGCCCAAGGACACTTTTACAAAATA 1248  
531 PheArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeu 549  
1249 TTTAAAGGTGTAAGAAGAGAAGTTGAGATTATGTCAGTCAGTCGACGAAACCGAAGTCTT 1308  
550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaIleSer 569  
1309 TTGAAGTCTTAGATAAAGCACAAGCACTATTTCAGAGTCTTCTTTGAAGCAGCAAGC 1368  
570 LeuMetSerGlnValSerProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589  
1369 ATGATGAGTCAGCTTCTCACAAGCAATTTGGTTTGAATTATGAGATGATGTCTGTGGA 1428  
590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608  
1429 GAGGAGAACAATTTGGTTTCAAGAGTTTGTAAATTTGGATCACTGGATACATACCTGAAG 1488  
609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628  
1489 AAGAACAAAATCTATAAATATATTATGGAACCTTGGAGTGGCAAGCAGTTGCATCG 1548  
629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648  
1549 GCCATGCATCTCTCGAAGAAAATCCCTTATTATCGGAATGTGTGCTTAAATAATC 1608  
649 LeuLeuAlaArgGluGlyGly---AspGlyAsnProPropheIleLysLeuSerAsp 666  
1609 CTGCTTATCAGAGAAGAAGACAGAGAACCGGGAACCCACCTTTCATCAAACTTAGTGAT 1668  
667 ProGlyValSerProThrValLeu-----SerLeuGluMetLeu 679  
1669 CCTGGCATTAGCATTTACAGTTCTACCGAAGGACATTTCTCTCTGTGTTCCTCAAGTCTT 1728  
680 ThrArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699  
1729 CAGGAGAATACCATGGTACCACTGATGTCATGAGATCCCTAAATACTTAACCTG 1788  
700 GluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAla 719  
1789 GCAACAGACAAGTGGAGCTTCGGGACCACTCTGTGGAGATCTCGAGTGGAGGAGATAAG 1848  
720 HistLeuThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeu 739  
1849 CCCCTGAGTCTCTGGATTCTCAAGAAAGCTGAGTTCTTATGAAGATAAGCATCAGCTT 1908  
740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759  
1909 CCTGCCCAAGTGGACAGATTGGCAACCTTATAAATAATTCATGAGTGGATGAGCCA 1968  
760 GlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779  
1969 GATTTCAGGCTGCTTTCAGAGCTGTCATCGGTGATCTTAAACAGCCTGTTTACTCCAGAT 2028  
780 TyrGluLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCys 798  
2029 TATGAACACTACTACAGAA---AATGACATGCTACCAACATGAGATAGTGGCTTAGGG 2085  
799 ValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeu 818  
2086 TTTTCTGGTCT-----TTTGAAGACAGGACCCCTACACAGTTTGAAGAGAGACACTT 2139  
819 LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp 838  
2140 AAGTTTCTACAGCAGCTTGGCAAGGTAATCTCGGAGTGTGGAGATGTGCCCTATGAC 2199  
839 ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858

QY 494 ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr 510  
Db 1129 GTTACAGCTCTCACCACCAATATTAGTGAATCAATGGTGTTCACCAA 1188  
QY 511 IleProThrAspSerLeuGluTriPheHisGluLeuGlyHisGlySerPheThrLysIle 530  
Db 1189 ATCAGGAATGAAGATTTCATATTATTAATGAAGCCTTGCCCAAGGACACTTTTACAAAATA 1248  
QY 531 PheArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeu 549  
Db 1249 TTTAAAGGTGTAAGAAGAGAAGTTGAGATTATGTCAGTCAGTCGACGAAACCGAAGTCTT 1308  
QY 550 LeuLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaIleSer 569  
Db 1309 TTGAAGTCTTAGATAAAGCACAAGCACTATTTCAGAGTCTTCTTTGAAGCAGCAAGC 1368  
QY 570 LeuMetSerGlnValSerProHisLeuValLeuLeuHisGlyValCysMetAlaGly 589  
Db 1369 ATGATGAGTCAGCTTCTCACAAGCAATTTGGTTTGAATTATGAGATGATGTCTGTGGA 1428  
QY 590 ---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArg 608  
Db 1429 GAGGAGAACAATTTGGTTTCAAGAGTTTGTAAATTTGGATCACTGGATACATACCTGAAG 1488  
QY 609 LysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyr 628  
Db 1489 AAGAACAAAATCTATAAATATATTATGGAACCTTGGAGTGGCAAGCAGTTGCATCG 1548  
QY 629 AlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysVal 648  
Db 1549 GCCATGCATCTCTCGAAGAAAATCCCTTATTATCGGAATGTGTGCTTAAATAATC 1608  
QY 649 LeuLeuAlaArgGluGlyGly---AspGlyAsnProPropheIleLysLeuSerAsp 666  
Db 1609 CTGCTTATCAGAGAAGAAGACAGAGAACCGGGAACCCACCTTTCATCAAACTTAGTGAT 1668  
QY 667 ProGlyValSerProThrValLeu-----SerLeuGluMetLeu 679  
Db 1669 CCTGGCATTAGCATTTACAGTTCTACCGAAGGACATTTCTCTCTGTGTTCCTCAAGTCTT 1728  
QY 680 ThrArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeu 699  
Db 1729 CAGGAGAATACCATGGTACCACTGATGTCATGAGATCCCTAAATACTTAACCTG 1788  
QY 700 GluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPheGlnArgGlyProAla 719  
Db 1789 GCAACAGACAAGTGGAGCTTCGGGACCACTCTGTGGAGATCTCGAGTGGAGGAGATAAG 1848  
QY 720 HistLeuThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeu 739  
Db 1849 CCCCTGAGTCTCTGGATTCTCAAGAAAGCTGAGTTCTTATGAAGATAAGCATCAGCTT 1908  
QY 740 ProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspPro 759  
Db 1909 CCTGCCCAAGTGGACAGATTGGCAACCTTATAAATAATTCATGAGTGGATGAGCCA 1968  
QY 760 GlyArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAsp 779  
Db 1969 GATTTCAGGCTGCTTTCAGAGCTGTCATCGGTGATCTTAAACAGCCTGTTTACTCCAGAT 2028  
QY 780 TyrGluLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCys 798  
Db 2029 TATGAACACTACTACAGAA---AATGACATGCTACCAACATGAGATAGTGGCTTAGGG 2085  
QY 799 ValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeu 818  
Db 2086 TTTTCTGGTCT-----TTTGAAGACAGGACCCCTACACAGTTTGAAGAGAGACACTT 2139  
QY 819 LysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAsp 838  
Db 2140 AAGTTTCTACAGCAGCTTGGCAAGGTAATCTCGGAGTGTGGAGATGTGCCCTATGAC 2199  
QY 839 ProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValPro 858





	Db	418	 GAGATTTTGCACCACTAATAACTCGAAACGGTGGAATTTCAGTGTGCTCAGAGGGAAA	477
	Qy	281	-----AsnAspGluLeuPheGlnThrPheCysAspPheProGlu	293
	Db	478	CATAANGNAAGTGACACTGACAGACGAGCTACAGTTATATGTGATTTCCCTCAT	537
	Qy	294	IleValaspValSerIleAsnGlnAlaProArgValGlyProAlaGlyLgLUHISArgLeu	313
	Db	538	ATTATTTGATGTTCAGTTATTAAACAAGCAAATCAG--GAATGCTCAACTGMAAGTAGATT	594
	Qy	314	ValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGlu	333
	Db	595	GTCACCCTGCCAACAGCAGCGGGAAGGCTCTTGAAATAGAACTTAGCTCATTAAGAANA	654
	Qy	334	AlaLeuSerPheValAlaLeuValAspGlyTyrrPheArgLeuIleCysAspSerArgHis	353
	Db	655	GCCTTGTCATTCGTGTCATTAATTGACGGGATTTACAGACTAACCTGCCGATGACACCAT	714
	Qy	354	TyrPheCysIysGluValAlaProProArgLeuLeuGluGluGluAlaAspValCysHis	373
	Db	715	TACCTCTGCAAGAGGTGGCTCCCCAGCTGTGTTCTGAGAACATACACGCAACTGCCAC	774
	Qy	374	GlyProIleThrLeuAspPheAlaIleHisIysLeuIysAlaLaGlyserLeuProgly	393
	Db	775	GGCCCAATTTCAATGGATTTTGGCCATCAGCAAACTTAAAGAAAGGCAGGAAACACGACTGGA	834
	Qy	394	ThrTyrtileLeuArgSergSerProGlnAspTyrrAspSerPheLeuLeuThrAlaCysVal	413
	Db	835	CTGTATGTACTTCGATGTAGCCCTPAAGCACTTCAACAATACTTCTCGACCTTTGCCGCTT	894
	Qy	414	GlnThrProLeuGlyProAspTyrrLyseGlyCysLeuIleArgGlnAspProSerGlyAla	433
	Db	895	GAGCGAAAAAATGTTATTGAATAATAACACTGTTTGATTACAAAGAAATGAGAATGGAGAG	954
	Qy	434	PheSerLeuValGlyLeuSerGlnProHiIsargSerLeuArgGluLeuAlaAlaCys	453
	Db	955	TACAACCTCAGTGGCACTTAAGAGAACTTTCAGTAGCTCTTAAGGACCTTTTGAATTGTCTAC	1014
	Qy	454	TrpAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrrLeuThrSerCysCysAlaPro	473
	Db	1015	CAGATGAAACTGTGCGCTCAGACAGTATCATCTTCAGTTCACCAATGTGTCCTCCA	1074
	Qy	474	ArgProlysGluIysSerAsnLeuIleValValArgArgGlyCysAsnProAlaproAla	493
	Db	1075	AAGCCGAAAGATAAATCAAACTTCTGTCTTCAGA----ACAAATGGTGTTTCTCAT	1128
	Qy	494	ProGlyCysSerProSerCysCysAla-----LeuThrGlnLeuSerPheHisThr	510
	Db	1129	GTTCCAGCTCTCACCAACATTACAGAGCATATAATGTGAATCAATATGGTGTTCACAAA	1188
	Qy	511	IleProthrAspSerLeuGluTyrrHisIleGluAsnLeuGlyHisGlySerPheThrLysIle	530
	Db	1189	ATCAGGAATCAAGATTGATATTATTAATGAAGCCTTGCCCAAGSCATTTTTACAAAAATA	1248
	Qy	531	PheArgGlySerArgArgGluValValAsp---GlyGluThrHisAspSerGluValLeu	549
	Db	1249	TTTAAAGGTGTAAGAGAGAAGATTGGAGATTATGTCAGCTGCACGAAACCGAAGTTCCT	1308
	Qy	550	LeuIysValMetAspSerArgHisArgAsnCysMetGlnSerPheLeuGluAlaAlaser	569
	Db	1309	TTGAAGTCTGATAAAGCACATAGAAACUATTATTCAGAGCTCTTCTTTGAAGCAGCAAGC	1368
	Qy	570	LeuMetSerGlnValSerTyrrProHisLeuValLeuLeuHisGlyValCysMetAlaGly	589
	Db	1369	ATGATGAGTCAGCTTCTCACAGCATTTGGTTTTGAAATTATGAGTATGTGTGTGGA	1428
	Qy	590	--AspSerIleMetValGlnGluPheValTyrrLeuGlyAlaIleAspMetTyrrLeuArg	608
	Db	1429	GAGGAGAACAATTTTGGTTCAAGAGTTTGTAAAATTTGGATCCTGGATACATACCTGAAG	1488
	Qy	609	LysArgGlyHisLeuValSerAlaSerTyrrPylsLeuGlnValThrIysGlnLeuAlaTyrr	628

QY 978 ValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLeuSerAspAsn 997  
 Db 2620 GTAAAGGAGCCAGGGGAAAGCCCAATATCTGTACGCACCTGAATCCTTGACGGAGAGC 2679  
 QY 998 IlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThr 1017  
 Db 2680 AAGTTTTCTGTGGCTCAGATGTGTGGAGCTTTGGAGTGGTTCTATACGAACCTTTTCACA 2739  
 QY 1018 TyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArg 1037  
 Db 2740 TACATCGAAGAGTAAAGTCCACCCGCGGAATTTATCGGAATGATTGGCAATGATAAA 2799  
 QY 1038 GluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGluGlyArgLeuPro 1056  
 Db 2800 CAAGGGCAAAATGATTGTGTGTTCCATTGTATAGAGCTACTGAGAGCAACGGAAAGATTGCCA 2859  
 QY 1057 ProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAlaProGlu 1076  
 Db 2860 AGGCCAGAGGATGCCAGATGAGATTTATGTGATCATGCACAGAGTGTCTGGAAACAACAT 2919  
 QY 1077 ProHisAspArgProAlaPheAlaThrLeuSer 1087  
 Db 2920 GTGAGCCAGCGTCCCTCCTTCAGGGACCTTTCC 2952

Search completed: February 4, 2004, 06:28:04  
 Job time : 295.373 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 06:28:12 ; Search time 1087.04 Seconds  
(without alignments)  
3724.149 Million cell updates/sec

Title: US-09-397-967A-16  
Perfect score: 5860  
Sequence: 1 MAPSEETPLIPORSCLSS.....RPAPATLSQLDPLMRGPG 1099

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delgap 6.0 , Delext 7.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US09397967/runat\_03022004\_175633\_28650/app\_query.fasta\_1.1486  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HRAPIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09397967@cgn1.1.383@runat\_03022004\_175633\_28650  
-NCPU=6 -ICPU=3 -NO.WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
-----			

1	4713.5	80.4	5118	13	US-09-814-353-19356	Sequence 19356, A
2	4597	78.4	3620	12	US-10-305-720-1055	Sequence 1055, Ap
3	2636.5	45.0	3731	12	US-10-191-803-200	Sequence 200, App
4	2594	44.3	4078	12	US-10-305-720-297	Sequence 297, App
5	2467	42.1	7126	12	US-10-062-674-2214	Sequence 2214, Ap
6	1927	32.9	5072	12	US-10-295-027-697	Sequence 697, App
7	1927	32.9	5073	15	US-10-084-817-230	Sequence 230, App
8	1908.5	32.6	4191	11	US-09-962-854A-1	Sequence 1, Appli
9	1904	32.5	3541	10	US-09-880-107-2379	Sequence 10, Appl
10	1842	31.4	3541	13	US-09-873-367C-163	Sequence 2379, Ap
11	1842	31.4	3538	11	US-09-962-854A-2	Sequence 2, Appli
12	1817.5	31.0	4080	12	US-10-305-720-1353	Sequence 1353, Ap
13	1446.5	24.7	5228	12	US-10-062-674-2215	Sequence 2215, Ap
14	1281.5	21.9	2597	13	US-10-205-219-2	Sequence 2, Appli
15	746	12.2	2327	12	US-10-131-410-22	Sequence 22, Appl
16	713	12.2	2338	9	US-09-925-302-90	Sequence 90, Appl
17	675.5	11.5	1179	15	US-10-211-088-131	Sequence 131, App
18	609.5	10.4	2027	10	US-09-771-161A-29	Sequence 29, Appl
19	483.5	8.3	975	11	US-09-962-854A-5	Sequence 5, Appli
20	482.5	8.2	753	9	US-09-910-943-636	Sequence 636, App
21	482	8.2	3845	13	US-10-007-926A-110	Sequence 110, App
22	482	8.2	3845	13	US-10-101-510-4	Sequence 4, Appli
23	482	8.2	3845	13	US-10-021-660-34	Sequence 34, Appl
24	482	8.2	3914	14	US-10-044-090-148	Sequence 148, App
25	475	8.1	3765	13	US-10-207-498-5	Sequence 5, Appli
26	475	8.1	3768	9	US-09-811-123-8	Sequence 8, Appli
27	475	8.1	3768	9	US-09-811-115-2	Sequence 2, Appli
28	475	8.1	4530	10	US-09-877-177-11	Sequence 11, Appl
29	475	8.1	4530	12	US-10-159-563-208	Sequence 208, App
30	475	8.1	4530	12	US-10-435-696-10	Sequence 10, Appl
31	475	8.1	4530	13	US-10-007-926A-119	Sequence 119, App
32	475	8.1	4530	13	US-10-101-510-124	Sequence 124, App
33	475	8.1	4530	13	US-10-338-730-1	Sequence 1, Appli
34	475	8.1	4530	13	US-10-116-275-131	Sequence 131, App
35	475	8.1	4530	13	US-10-272-437A-27	Sequence 27, Appl
36	475	8.1	4530	13	US-10-117-937-595	Sequence 595, App
37	475	8.1	4530	15	US-10-177-293-125	Sequence 125, App
38	475	8.1	9274	9	US-09-811-123-7	Sequence 7, Appli
39	475	8.1	9274	9	US-09-811-115-1	Sequence 1, Appli
40	474	8.1	3768	10	US-09-854-356-9	Sequence 9, Appli
41	474	8.1	3768	10	US-09-930-125-1	Sequence 1, Appli
42	474	8.1	3768	13	US-10-313-644-1	Sequence 1, Appli
43	474	8.1	4473	11	US-09-441-411-5	Sequence 5, Appli
44	474	8.1	4473	13	US-10-101-510-81	Sequence 81, Appl
45	474	8.1	4473	15	US-10-146-473-32	Sequence 32, Appl

# ALIGNMENTS

RESULT\_1  
US-09-814-353-19356  
; Sequence 19356, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Lillie, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF OVARIAN CANCER  
; FILE REFERENCE: MRI-006B  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25

;; PRIOR APPLICATION NUMBER: US 60/257,672  
;; PRIOR FILING DATE: 2000-12-21  
;; NUMBER OF SEQ ID NOS: 22037  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 19356  
;; LENGTH: 5118  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc feature  
;; LOCATION: 1, 2, 3, 4, 5114, 5115, 5116, 5117, 5118  
;; OTHER INFORMATION: n = A,T,C or G

US-09-814-353-19356

Alignment Scores:  
Pred. No.: 0 Length: 5118  
Score: 4713.50 Matches: 912  
Percent Similarity: 88.17% Conservative: 64  
Best Local Similarity: 82.38% Mismatches: 122  
Query Match: 80.44% Indels: 10  
DB: 13 Gaps: 5

US-09-397-967A-16 (1-1099) x US-09-814-353-19356 (1-5118)

QY 1 MetAlaProSerGluGluThrProLeuLeuProGlnArgSerCysSerLeuSerSer 20  
DB 96 ATGGACCTCCCAAGTGAAGACGCGCCCTGATCCCTCAGCGTTTCAGCGCTCTGTGCC 155  
QY 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
DB 156 ACGGAGCTGGTCCCTGCATGTGCTGCCCGCTCGGGGCCCGGGCCCGCCAGCGC 215  
QY 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
DB 216 CTATCTTCTCTTGGGACCACTTGGCTGAGGACCTGTGGTGAGCGTGCAGCGTCCAAAGGCC 275  
QY 61 CysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
DB 276 AGCGGACATCTGCTGTATCCACTCCCTCTTGTCTGTGGCCAGGAGACCTGTCTGC 335  
QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
DB 336 TGGTTCCTCCCGAGCCACATCTTCCTGGTGGAGATGCCAGCACCCCAAGTCTGTGTAC 395  
QY 101 ArgLeuArgPheTyrPheProAspTyrPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
DB 396 AGGATTTCGCTTTTACTTCCCAATTGGTTTGGGCTGGAGAAGTGCCACCGCTTCGGGCTA 455  
QY 121 ArgGlyAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
DB 456 CGCAAGATTGGCCAGTGCTATCTTGTACCTGCCAGTCTCTGGAGCACCTCTTTGCCAG 515  
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGly 160  
DB 516 CACCGCAGTGACCTGGTGGTGGCGGCTCCCGTGGGCTCAGTCTCAAGGAGCAGCGGT 575  
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
DB 576 GAGTGTCTACGCTGGCGCGTGTGGACCTGGCCGGATGGCGGAGAGCGAGGCCACGCG 635  
QY 181 ProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAsp 200  
DB 636 CCGGGAGAGCTGTGAAGATGTCTCAGTACAGGCTGCTACCCCCAAGCTTGGCGGAC 695  
QY 201 ValIleGlnGlyGlnAspPheValThrArgArgArgIleArgThrValValLeuAla 220  
DB 696 CTGATCCAGGCGCTGAGCTTCGTACCGGAGGCGGTATTCGGAGGACGCTGCCAGAGCC 755  
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaLysTyrIleLe 240  
DB 756 CTGGC-CCGCGTGGCCCGCTGCCAGGAGACCGGACCTCGCTCATGCGCAAGTACATCAT 814  
QY 240 uAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGln 260

DB 815 GGACCTGGAGCGGTGGATCCAGCGGGCGCGCCAGACCTTCCACGTGGCGCTCCCTGG 874  
QY 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277  
DB 875 GGCCCTTGGTGGCCACAGCAGCGGCTGGGCTGCTCCGCGTGGCTGTGTGACGGCGGATCGC 934  
QY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluLeuValAs 296  
DB 935 CTGGACCCAGGAGAACAGAGGTCTCCAGCCCTTCTCGGACTTTCAGAAATCGTAGA 994  
QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316  
DB 995 CATTAGCATCAAGCAGAGCGCCCGCGGCTGGCCCGCGGAGAGACCGCGCTCTCTCTG 1054  
QY 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336  
DB 1055 TACCAGGACAGACAAACAGATTTTAGAGGCCGAGTTCCAGGGCTGCGCCAGGCTCTGTC 1114  
QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356  
DB 1115 GTTCGTGGCGCTCTGTGGACGCTACTTCGGGCTGACACCGGACTCCACGACTTCTCTG 1174  
QY 356 sLysGluValAlaProProArgLeuLeuGluGluAlaAspValCysGlyProI 376  
DB 1175 CAAGGAGGTGGCACCGCGAGGCTGTGGAGGAAGTGGCCGAGCAGTGCACGGCCCCAT 1234  
QY 376 eThrLeuAspPheAlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIl 396  
DB 1235 CACTCTGGAGCTTTGCCATCAACAGCTCAAGACTGGGGGCTCACGTCTGGCTCTTATGT 1294  
QY 396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416  
DB 1295 TCTCCGCGCAGCCCCCAGGACTTTGACAGCTTCTCTCTCTCTCTGTGTGTCAGAACCC 1354  
QY 416 oLeuGlyProAspTyrLysGlyCysLeuLeuArgGlnAspProSerGlyAlaPheSerLe 436  
DB 1355 CCTTGGTCTGTATTAAAGGCTGCTCATCGCGCGCAGCCCCACAGGAACCTTCTTCT 1414  
QY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAalaCysTrpAsnSe 456  
DB 1415 GGTGGCTCTCAGCGCCACCCACAGCAGTCTTTCAGAGCTCTCTGGCAACCTGTGGGATGG 1474  
QY 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLy 476  
DB 1475 GGGGCTGCAGTAGATGGGTGGCAGTCCCTCACTTCTCTGTATCCACAGACCCAA 1534  
QY 476 sGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGlyCy 496  
DB 1535 AGAAAGTCCAACTGATCGTGTCCAGAGAGTCCAGCCCAACCCACATCATCTCTTGGT 1594  
QY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
DB 1595 TCAGGCCCAATCCCAATACAGCTGAGTCAGATGACATTCACAAGATCCCTCTGACAG 1654  
QY 515 rLeuGluTyrHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535  
DB 1655 CTTGGAGTGGCATGAGAACCTGGGCCATGGTCTCTTACCACAGATTATCCGGGCTGTCTG 1714  
QY 535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555  
DB 1715 CCATGAGTGGTGGATGGGAGGCCCGAAAGACAGAGAGTCTCTGAAGTCTATGATGTC 1774  
QY 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575  
DB 1775 CAAGACCAAGACTCATGAGTCAATCTCTGGAGGACGAGCTGTGATGAGCAAGTGTG 1834  
QY 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595  
DB 1835 GTACCGGCACTCTGTGTCTGCCAGCGGTGTGATGGTGGAGACAGCAGCAGTGGTGA 1894  
QY 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615

Db 1895 GGAATTTTGTACACCTGGGGCCATAGACATGATCTCGGAAAAAGTGGCCACCTGGTGGC 1954  
Qy rAlaserTrrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAenTyrLeuGluAs 635  
Db 1955 AGCCAGCTGGAGCTGCGAGGTGCTCAACACAGCTGACCTTACCCCTCAACTATCTGGAGGA 2014  
Qy pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyG1 655  
Db 2015 CAAGAGCCTGCCCATGCAATGTCTCTGCCCGAAGGTGCTCCTGGCTCGGAGGGGGC 2074  
Qy yAspGlyAsnProProPheIleLysLeuSerAspProLysValSerProThrValLeuSe 675  
Db 2075 TGATGGGAGCCCGCCCTTCATCAAGAGTGAGTGACCTCGGGGTCAAGCTGCTGTTAAG 2134  
Qy rLeuGluMetLeuThrAspArgIleProTrrpValAlaProGluCysLeuGlnGluAlaG1 695  
Db 2135 CCTGGAGATGCTCACCGACAGATCCCTGGGTGGCCCCCGAGGTCTCCGGGAGGGCGCA 2194  
Qy nThrLeuCysLeuGluAlaAspLysTrrpGlyPheGlyAlaThrThrTrrpGluValPheG1 715  
Db 2195 GACACTTAGCTTGAAGCTGACAAAGTGGGGCTTCGGCGCCACAGCTCGGGAGGTGTTAG 2254  
Qy nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLysPheTyrGluAs 735  
Db 2255 TGGCGTCAACATGCCCATCAGTGGCCCTGGATCTCTGAAGAAACTCCAAATTTTATGAGA 2314  
Qy pGlnGlyGlnLeuProAlaLeuLysTrrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755  
Db 2315 CCGCAGCAGCTGCGGCCCCCAAGTGGACAGAGCTGGCCCTGCTGATTCAACAGTGCAT 2374  
Qy tAlaTyrAspProGlyArgProSerPheArgAlaIleLeuArgAspLeuAenGlyLe 775  
Db 2375 GGCCTATGAGCGCGTCCAGAGGCCCTCTCTCCGAGCGCTCATTCTGACCTCAATAGCCT 2434  
Qy vIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795  
Db 2435 CATCTCTTCAGACTATGAGTCTCTCTCAGACCCACACACCTGGTGGCCCTCGACCTCGTGA 2494  
Qy pGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluG1 815  
Db 2495 TGGGCTGTGG---AATGGTCCCGAGCTCTATGCTGCCAAGACCCACAGCTCTTCGAGGA 2551  
Qy uArgHisLeuLysTrrpIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCy 835  
Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGCGAAGGCACTTTGGCAGCTGGAGCTGTG 2611  
Qy sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHi 855  
Db 2612 CCGCTATGACCCCGCTAGCGGCAATACAGAGTGGCTGTGGCCGTGAACAGCTGCAGCA 2671  
Qy sSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHi 875  
Db 2672 CAGCGGCCAGACGACGAGAGGACTTTCAGCGGGAGATTTCAGATCTCAAGACACTGCA 2731  
Qy sSerAspPheIleValLysTrrpArgGlyValSerTyrGlyProGlyArg-GlnSerLeuA 895  
Db 2732 CAGTGAATTCATTGTCAAGTATCGTGTGTGTCAGTATGGCCCGCGCGCCAGAGCTGCG 2791  
Qy rGluValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgG 915  
Db 2792 GGCTGGTGCATGGAGTACCTGCCAGCGCTGCTTGGCGGACTTCTCTGCGCGCCAGCCGCG 2851  
Qy lY---LeuHisThrAspArgLeuLeuLeuPheAlaTrrpGlnIleCysLysGlyMetGluT 934  
Db 2852 CGGCGCTCGATGCCAGCGCGCTCTCTCTATTCTTCGAGATCTGCAAGGGATGGAGT 2911  
Qy YrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluS 954  
Db 2912 ACCTGGGCTCCCGCGCTGGGTGCGACCGGACCTGGCGCGCGCGGCAACATCTCTGTGAGA 2971  
Qy eGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysA 974  
Db 2972 GCGAGGCGACAGTACGCTGACTTCGGCCTAGCTTAAGCTGCTGCCGCTTGACAAAG 3031

Qy 974 spTyrTyrValValArgGluProGlyGlnSerProIlePheTrrpTrrpAlaProGluSerL 994  
Db 3032 ACTACTACGTGGTCCGGAGCAGGCGCAGAGCCCATTTTCTGGTATGCCCCCGAATCCC 3091  
Qy eUserAspAsnIlePheSerArgGlnSerAspValTrrpSerPheGlyValValLeuTyrG 1014  
Db 3092 TCTCGGACAAACATCTTCTCTGCCAGTCAGACGCTCGAGGCTTCGGGGTCTCTGTACG 3151  
Qy luLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetG 1034  
Db 3152 AGCTCTTCACCTACTGCGACAAAAGCTGCAGCCCTCGGCCGAGTCTCTCGCGATGATGG 3211  
Qy lyProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuLeuAlaGluGlyArgA 1054  
Db 3212 GATGTAGCGGAGTGTCCCGCCCTCTGCCGCTCTTGGAACTCTCGAGAGGGGCCAGA 3271  
Qy rGLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpA 1074  
Db 3272 GGCTGCCGCGCCTCTCTGCCCTGCTGAGGTTTCAGAGCTCATGAGCTGTGCTGGG 3331  
Qy laProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuT 1094  
Db 3332 CCCCTAGCCACAGACCGGCAATCATTCAGCGCCCTGGGCCCCAGCTGGACATGCTGT 3391  
Qy rPArgGlyArgProGly 1099  
Db 3392 GGAGCGAAGCCGGGG 3408  
RESULT 2  
US-10-305-720-1055  
; Sequence 1055, Application US/10305720  
; Publication No. US20040010136A1  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.  
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression  
; FILE REFERENCE: PA-0002-1 CON  
; CURRENT APPLICATION NUMBER: US/10/305,720  
; PRIOR FILING DATE: 2002-11-26  
; PRIOR APPLICATION NUMBER: 09/016,434  
; PRIOR FILING DATE: 1998-01-30  
; NUMBER OF SEQ ID NOS: 1490  
; SOFTWARE: PERL Program  
; SEQ ID NO 1055  
; LENGTH: 3620  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: GenBank ID No. US20040010136A1 g1039418  
US-10-305-720-1055  
Alignment Scores:  
Pred. No.: 0 Length: 3620  
Score: 4597.00 Matches: 895  
Percent Similarity: 86.80% Conservatives: 65  
Best Local Similarity: 80.92% Mismatches: 123  
Query Match: 78.45% Indels: 24  
DB: 12 Gaps: 7  
US-09-397-967A-16 (1-1099) x US-10-305-720-1055 (1-3620)  
Qy 1 MetAlaProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSer 20  
Db 96 ATGGCACTTCAAGTGAAGAGAGCGCCCTGATCCCTCAGCGTTTCATGCAGCTCTTGTGCC 155  
Qy 21 SerGluAlaGlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArg 40  
Db 156 ACGAGAGCTGTGTCCTGTCATGTGCTGCTGCCGCTCGGGCGCCCGGGCCCCCGAGCGC 215  
Qy 41 LeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla 60  
Db 216 CTATCTTCTCTTTGGGAGACCACTTGGCTGAGGACCTGTGCTGAGCTGCCAGGCTGCCAAGGCC 275

QY 61 CysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCys 80  
DB 276 AGCGGCAATCCCTGCTGTGTACACATCCCTCTTTGTCTGGCCACCGAGACCTGTCCTGC 335  
QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
DB 336 TGGTTCCCGCCGAGCCACATCTTCCCGTGGAGGATGCCAGACCCCAAGTCTCTGTGTAC 395  
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeuGluThrCysHisArgPheGlyLeu 120  
DB 396 AGGATTCCGCTTTACTTCCCAATTGGTTGGGCTGGAGAGTGCCACCGCTTCGGGCTA 455  
QY 121 ArgIysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGln 140  
DB 456 CCACAGGATTTGGCCAGTGCTATCTTGACCTGCCAGTCTCTGGAGCACCTCTTTGCCAG 515  
QY 141 HisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIysGluGlnGly 160  
DB 516 CACCGCAGTGACTGGTGGTGGCGGCTCCCGTGGGCTCAGTCTCAAGGAGCAGGCT 575  
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaArg 180  
DB 576 GAGTGTCTCAGCTCGCCCTGTGGACCTGGCCCGGATGGCGAGAGAGGCCAGCGG 635  
QY 181 ProGlyGluLeuLeuIysThrValSerTyrIysAlaCysLeuProProSerLeuArgAsp 200  
DB 636 CCGGAGAGCTGCTGAAGACTGTACGTACAAGGCTGCTTACCCCAAGCTTGCAGCAG 695  
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgArgIleArgArgThrValValLeuAla 220  
DB 696 CTGATCCAGGGCTGAGCTTCTGAGCGCGAGCGGTATTCGGAGGACGCTGCGCAGAGCC 755  
QY 221 LeuLeuProCysGlyArgLeuProGlyArgPro-TyrAlaLeuMetAlaIysTyrIleLeu 240  
DB 756 CTGCG-CCGCGTGGCGGCTGCGCAGCAGACCGGCACTCGCTCATGGCCCAAGTACATCAT 814  
QY 240 uAspLeuGluArgCysLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProG 260  
DB 815 GGACCTGGAGCGCTGGATCCAGCGCGGCGCGGAGACCTTCCAGCTGGCGGCTCCCTGG 874  
QY 260 yAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAspAsnGlyIlePr 277  
DB 875 GGCCTTGTGTGCCACGACGGGCTGGGCTGCTCCGCGTGGTGGTGACGGCGCATCGC 934  
QY 277 oTrpSerSerAsnAsp---GluLeuPheGlnThrPheCysAspPheProGluIleValAs 296  
DB 935 CTGGACCCAGGAGAACAGGAGGCTCTCCAGGCTTCTGCGACTTTCCAGAAATCGTAGA 994  
QY 296 pValSerIleAsnGlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrVa 316  
DB 995 CATTAGCATCAAGAGGCGCGCGCTGGCCCGCGGAGACACCGCTGCTGCTACTGT 1054  
QY 316 lThrArgMetAspGlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSe 336  
DB 1055 TACCAGGACAGACAACAGATTTTAGAGGCGGAGTTCCAGGCGTGCCCGAGGCTCTGTC 1114  
QY 336 rPheValAlaLeuValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCy 356  
DB 1115 GTTCGTGGCGCTCGTGAGCGGCTACTTCCGGCTGACCCAGGACTCCCGAGCACTTCTCTG 1174  
QY 356 sIysGluValAlaProProArgLeuLeuGluGluAlaAspValCysHisGlyProI 376  
DB 1175 CAAAGAGGTGGCACCGCGAGGCTGTCTGGAGGAGTGGCCGAGAGTGGCCAGGCTCCAGCC 1234  
QY 376 eThrLeuAspPheAlaIleHisIysLeuLysAlaAlaGlySerLeuProGlyThrTyrI 396  
DB 1235 CACTCTGGACTTTGCCATCAACAAGCTCAAGACTGGGGGCTCAGCTCTGGCTCTATGT 1294  
QY 396 eLeuArgArgSerProGlnAspTyrAspSerPheLeuLeuThrAlaCysValGlnThrPr 416  
DB 1295 TCTCCGCGGACCGCCAGGACTTTGACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1354

QY 416 oLeuGlyProAspTyrIysGlyCysLeuIleArgGlnAspProSerGlyAlaPheSerIle 436  
DB 1355 CTTGGTCTCTGATTATAAGGCTGCTCATCGGCGCAGCCACAGGAACCTCTCTCTCT 1414  
QY 436 uValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSe 456  
DB 1415 GGTGGCTCAGCCAGCCACAGCAGTCTTCGAGAGCTCCGGAACCTCTCTGGGATGG 1474  
QY 456 rGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgProly 476  
DB 1475 GGGGCTGCAGTAGATGGGTGGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1534  
QY 476 sGlyIysSerAsnLeuIleValValArgGlyCysAsnProAlaProAlaProGlyCy 496  
DB 1535 AGAAAGTCCAACTGATCGTGTCCAGAGAGTCCAGCCACCCACATCATCTCTGTGT 1594  
QY 496 sSerPro---SerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSe 515  
DB 1595 TCAGCCCAATCCCAATACCACTGAGTCAGATGACATTTCAAGATCCCTGTGTGACAG 1654  
QY 515 rLeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerAr 535  
DB 1655 CTGTGAGTGGCATGAGACCTGGGCTCTTCCACCAAGATTTACCGGGGCTGTCTG 1714  
QY 535 gArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAspSe 555  
DB 1715 CCATGAGTGTGGTGGGAGGCGCGMAAGACAGAGGTGCTGCTGAAGTCTATGATGC 1774  
QY 555 rArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSe 575  
DB 1775 CAAGCACAGAACTGCTGAGTGCATATCTCTGGAACAGCAGGCTTGATGAGCAAGTGC 1834  
QY 575 rTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetValGl 595  
DB 1835 GTACCGGCTATCTGTGCTGCTCCAGCGCTGTGCTGCTGGAGACAGCAGCATGTCGA 1894  
QY 595 nGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSe 615  
DB 1895 GGAAATTTGTACCTTGGGCGCATAGACATATATCTGCGAAACAGTGGGCGACCTGTG 1954  
QY 615 rAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAs 635  
DB 1955 AGCCAGCTGGAGCTGAGTGGTCAACAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2014  
QY 635 pLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGl 655  
DB 2015 CAAAGGCTTCCCGCATGCAATGCTCTGCGCGAAGGTGCTCTGCTGCTGCTGCTGCTG 2074  
QY 655 yAspGlyAsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSe 675  
DB 2075 TGATGGGAGCGCCCTTTCATCAAGCTGAGTGACCTGGGGTTCAGCCCGCTGTGTTAG 2134  
QY 675 rLeuGluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGl 695  
DB 2135 CTTGGAGATGCTCACCGACAGATCCCTGGGTGGCCCGAGTGTCTCGGAGGCGCA 2194  
QY 695 nThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGl 715  
DB 2195 GACACTTAGCTTGGAGCTGACAACTGGGCTTCGGCGCCACCGTCTGGGAAGTGTAG 2254  
QY 715 nArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAs 735  
DB 2255 TGCGCTCACCATGCCCTCAGTCCCTGATCTCTGCTAAGAACTCAATTTTATGAGGA 2314  
QY 735 pGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMe 755  
DB 2315 CCGGACAGAGCTCCCGGCGCCCAAGTGGAGCAGAGCTGGGCTCTGCTGATTCACAGTGCAT 2374  
QY 755 tAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLe 775  
DB 2375 GGCCTATGAGCGGTCCAGAGGCGCTCTTCCGAGCGCTCATTCGTGCTCACTCAATAGCT 2434  
QY 775 uIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAs 795

```
Db 2435 CATCTCTTCAGACTATGAGTCTCTCTCAGACCCACACCTGGTGCCTGGCACCTCGTGA 2494
Qy 795 pgluleucysvalalaglyalaglnleuTyralaCysGlnAspProAlaIlePheGluGl 815
Db 2495 TGGGCTGTGG--AATGGTGCCAGCTCTATGCCCTGCCAAGACCCACAGATCTTCGAGA 2551
Qy 815 uArgHisLeuIstYrIleSerLeuLeuGlyIstYrIlePheGlySerValGluLeuCy 835
Db 2552 GAGACACCTCAAGTACATCTCACAGCTGGGCAAGGCAACTTTGGCAGGTGGAGCTGTG 2611
Qy 835 sArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValIstYrIleGlnHis 855
Db 2612 CCGTATGACCGCTAGGCGCAATACAGGTGCCCTGGTGGCGGAAACAGCTCCAGCA 2671
Qy 855 sSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnIleLeuLeuAlaLeuHi 875
Db 2672 CAGCGGCCAGACAGGAGGAGCTTTTCAGCGGAGATTTCAGATCCTCAAGCACTGCA 2731
Qy 875 sSerAspPheIleValIstYrIleArgGlyValSerTyrGlyProGlyArgGlnSerLeuAr 895
Db 2732 CAGTGATTTTCATTGCAAGTATCGTGTGTGAGTATGGCCGGGCCCGCCAGAGCTGG 2791
Qy 895 gLeuValMetGluTyrIleuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGl 915
Db 2792 GCTGCTCATGTGAGTACCTGCCAGCGGCTGCTTGGCGACTTCCTGCAGCGGCACCGCGC 2851
Qy 915 Y---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysIstYrIleMetGluTy 934
Db 2852 CGGCTCGATGCCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2911
Qy 934 rLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSe 954
Db 2912 CCTGGCTCCCGCGCTCGGTGACCGGACCTGGCGCCCGCAACATCTCTCTGTGGAG 2971
Qy 954 rGluAlaHisValIstYrIleAlaAspPheGlyLeuAlaIstYrIleLeuProLeuGlyLysAs 974
Db 2972 CGAGGCACAGCTCAAGATCGTACTCGGCTAGCTAAGCTGCTGGCGCTTGCAAGA 3031
Qy 974 pTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGluSerLe 994
Db 3032 CTACTAGTGTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 3091
Qy 994 uSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGl 1014
Db 3092 CTGGCAACAATCTCTCTCGCAGTCTGAGCTCTGGAGCTTCGGGTCTGCTGTAGCA 3151
Qy 1014 uLeuPheThrTyrCysAspIstYrIleSerProSerAlaGluPheLeuArgMetMetGl 1034
Db 3152 GCTCTTCACCTACTCGACAAAGCTGAGCGGCTCGGCGGAGTTCCTGGGATGATGG 3211
Qy 1034 YProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgAr 1054
Db 3212 ATGTAGCGGGAGTGTCCCGGCTCTCGGCGCTCTGGAACTGCTGGAGGAGGCGCAG 3271
Qy 1054 gLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAl 1074
Db 3272 GCTGGCGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3305
Qy 1074 aProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeuTr 1094
Db 3306 -----AGCGCGGCGGCTGAGCTGAGTTCAGGCTCTGAGAT-----TG 3346
Qy 1094 pArgGlyArgProGly 1099
Db 3347 GGCGGGGCTCGGGC 3362
```

## RESULT 3

US-10-191-803-200  
; Sequence 200, Application US/10191803  
; Publication No. US20040014040A1  
; GENERAL INFORMATION:  
; APPLICANT: MENDRICK, Donna

APPLICANT: PORTER, Mark  
APPLICANT: JOHNSON, Kory  
APPLICANT: HIGGS, Brandon  
APPLICANT: CASTLE, Arthur  
APPLICANT: ELASHOFF, Michael  
TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling  
FILE REFERENCE: 44921-5090US  
CURRENT APPLICATION NUMBER: US/10/191,803  
CURRENT FILING DATE: 2002-07-10  
PRIOR APPLICATION NUMBER: US 60/303,819  
PRIOR FILING DATE: 2001-07-10  
PRIOR APPLICATION NUMBER: US 60/305,623  
PRIOR FILING DATE: 2001-07-17  
PRIOR APPLICATION NUMBER: US 60/369,351  
PRIOR FILING DATE: 2002-04-03  
PRIOR APPLICATION NUMBER: US 60/377,611  
PRIOR FILING DATE: 2002-05-06  
NUMBER OF SEQ ID NOS: 1140  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 200  
LENGTH: 3731  
TYPE: DNA  
ORGANISM: Rattus norvegicus  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM\_031514

US-10-191-803-200

Alignment Scores:  
Pred. No.: 1,36e-261 Length: 3731  
Score: 2636.50 Matches: 538  
Percent Similarity: 65.77% Conservative: 196  
Best Local Similarity: 48.21% Mismatches: 349  
Query Match: 44.99% Indels: 33  
DB: 12 Gaps: 15

US-09-397-967A-16 (1-1099) x US-10-191-803-200 (1-3731)

```
Qy 4 ProSerGluGluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAla 23
Db 208 CCTGCACATCAGATGATGTTGCGAAGAAATGCTAACTCTGTGAAGCAGACAGAGCCA 267
Qy 24 GlyAlaLeuHisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPhe 43
Db 268 GTCCCTCAAGTGATACCTGTGACCATTCTCTGGGCAAGCTGAAGAGACTATCTGAAGTTT 327
Qy 44 SerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAlaCysGlyIle 63
Db 328 CCAATCGAGAGATGTTGCGAAGAAATTTGTGTGCTGCTCTTAAAGCATGTGGTATT 387
Qy 64 LeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPheSerCysTrpPhePro 83
Db 388 ACGCTGTGTATCATATATATGTTTGTGTTTAAATGAGTGAAGCAAGAGATCTGTATCCA 447
Qy 84 ProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuArg 103
Db 448 CCCAATCATGCTCTCCACATAGATGATCAACAGGCAATATATATCTCTACAGAAATAGG 507
Qy 104 PheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisArgPheGlyLeuArg 121
Db 508 TTTTACTTCTCATTTGTTACTGCAGTGGCAGCAACAGAACCTTATAGGTACGGAGTATCT 567
Qy 122 LysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
Db 568 CGTGGGCGCTGAAGCTCTCTCTGTGACGATTTGTGTCATGTTACCTTTTGTGTCAGTGG 627
Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGlnGlyGlu 161
Db 628 CGGCATGATTTTGTTCATGGGTGGAATAAGTACCTGTGACTCACCAGAACTCAGGAAGAA 687
Qy 162 PheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgPro 181
Db 688 TGTCCTTGGCATGCGAGTGTGTAGCATGTAGATAGTAAAGTAAAGGAGAGGAGGAGAGCTCCC 747
```

```
QY 182 GlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSerLeuArgAspVal 201
Db 748 CTGGCTGCTATTAACCTCATCAGCTACAAAGAGCTTCTTACCGAAGTGGCGGCGAAG 807
QY 202 IleGlnGlyGlnAsnPheValThrArgArgArgIleArgThrValValLeuAlaLeu 221
Db 808 ATCCAGGACTATCACATTTTAACCTCGAAGCGGATAAGGTACAGATTCCGCGAGTTCAAT 867
QY 222 LeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAsp 241
Db 868 CAGCAGTTTCAGTCAGTCAAGGCCACCTGCACAGAACTTAAAGTATCTCAATAAC 927
QY 242 LeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgVal 256
Db 928 CTGGAGACCTCGAGTCTCCCTTTATACACAGAACAGTTTGAAGTAAAGAACTCTGCAAGA 987
QY 257 GlyLeuProGlyAlaLeuGlnGluProGlyLeuLeuArgValAlaGlyAspGlnGlyIle 276
Db 988 GGT---CCCTTCAGGTGAGGAGATTTTGGCAACCATTAATAACTGGAACGCGTGAAT 1044
QY 277 ProTrpSerSer-----AsnAspGluLeuPheGlnThr 287
Db 1045 CAGTGTCTCAAGAGGAAACATTAAGAAAGTGAACACTCACAGAACAGGATTTTACAGTTG 1104
QY 288 PheCysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyPro 307
Db 1105 TATTTGTGATTTTCTCATATTATGATGTCAGTATTAAAGCAGCAATCAG---GAATGC 1161
QY 308 AlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGlu 327
Db 1162 TCAACTGAAAGTAGAGTTGTGACCGTCCACAGCAGGCGGAAGGCTTGGAAATAGAA 1221
QY 328 PheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeu 347
Db 1222 CTAGTCTCATTAAGAAAGCCTTGTCTATTGTGTCGTTAATGACGGGTATTACAGACTA 1281
QY 348 IleCysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGlu 367
Db 1282 ACTGCGGATGCACACCATTCCTCTGTAAAGAGTGGCTCTCCAGCCGTGCTTGAATC 1341
QY 368 GluAlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAla 397
Db 1342 ATACACAGCAACTGTACGCGCCCAATTTCAATGAGATTTTGCCATCAGCAAACTAAAGAAG 1401
QY 388 AlaGlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPhe 407
Db 1402 GCAGGAAACACAGACTGGACTGTATGTACTTCGATGTAGCCCTAAGGACTTCAACAAATAC 1461
QY 408 LeuLeuThrAlaCysValGlnThrProLeuGlyProAspTyrLysGlyCysLeuIleArg 427
Db 1462 TTCCTGACCTTTCGGTTTGAGCGAGAAATGTTATTGAAATATAACACACTGTTGATTACA 1521
QY 428 GlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnProHisArgSerLeuArg 447
Db 1522 AAGAATGAGAATGGAGAGTCAACCTCAGTGGGACTAAGAGGAACCTTCAGTAGTCTTAAG 1581
QY 448 GluLeuLeuAlaAlaCysTyrPheAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeu 467
Db 1582 GACCTTTTGAATGCTACCATGATGGAACCTGTGCGCTCAGACAGATATCATCTTCCAGTTC 1641
QY 468 ThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIleValValArgArgGly 487
Db 1642 ACCAATGCTGTCTCCAAAGCCGAAAGATAAATCAACCTTCTTGTCTTCAGA----- 1695
QY 488 CysAsnProAlaProAlaProGlyCysSerProSerCysCysAla-----LeuThr 504
Db 1696 ACAAAATGGTGTCTTCGATGTTTCACTCTCCACCAACATTACAGAGCCGCAATAAATGTGAAT 1755
QY 505 GlnLeuSerPheHisThrIleProThrAspSerLeuGluThrPheHisGluAsnLeuGlyHis 524
Db 1756 CAAATGGTGTTCACAAATCAGGAATGAAGATTTGATATTATTAATGAAGCCCTTGGCCAA 1815
QY 525 GlySerPheThrLysIlePheArgGlySerArgArgGluValValAsp---GlyGluThr 543
1816 GGCACCTTTTACAAAATATTTAAAGCTGTAAAGAGAGATTGGAGATTATGGTCAGCTG 1875
544 HisAspSerGluValLeuLeuLysValMetAspSerArgHisArgAsnCysMetGluSer 563
1876 CACGAAACCGAAGTCTCTTTGAAAGTCTTAGTAAGACACATAGAACTATTTCAGAGTCT 1935
564 PheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuValLeuLeuHis 583
1936 TTCCTTTGAAGCAGCAGCATGATGATGAGCTTTCTCACAAAGCTTTGGTTTGAATTAT 1995
584 GlyValCysMetAlaGly---AspSerIleMetValGlnGluPheValTyrLeuGlyAla 602
1996 GGAGTATGTCTGTGGAGAGGAGAACATTTTGGTTCAGAGCTTTGTAAGATTTGTAATAT 2055
603 IleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLysLeuGlnVal 622
2056 CTGGATACATACCTGAAAAAAACAAAATTCATTAATATATCTATGGAACTTGGAGTG 2115
623 ThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuProHisGlyAsn 642
2116 GCCAAACAGTTGGCATGGGCCATGCATTCCTCGAAGAAAAATCCCTTATTCATGGAAT 2175
643 ValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGlyAsnProPro 660
2176 GTGTGTGCCAAAATATCTCTCTTATCAGAGAAGAACAGAGAAACACAGGGAATCCACCT 2235
661 PheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGluMetLeuThr 680
2236 TTCATCAAACTTAGTGATCTCTGGGATTAGCATCACAGTTCTACCAAGAGACATCTTCAG 2295
681 AspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThrLeuCysLeuGlu 700
2296 GAGAGTAATACCTGGGTACCCTGAGTGCATTTGAAAAACCCCTAAAAATCTAACTTTGGCA 2355
701 AlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGlyProAlaHis 720
2356 ACAGACAGTGGAGCTTTGGCACCACTCTGTGGGAGATCTGCAGTGGAGAGAGCAAGCCT 2415
721 IleThrSerLeuProAlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuPro 740
2416 CTGAGTGTCTGAGCTCTCAAGAAAGCTGCAGTTCTATCAAGATAAGCACCAAGCTCCCA 2475
741 AlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGly 760
2476 GCACCCAGTGGACAGAGTTGGCAACCTTATAAATACTTGCATGGATTATGAGCCAGAC 2535
761 ArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyr 780
2536 TTCAGGCTGCTTCAGAGCTGTATCCGCTGATCTGAACAGCCTGTTTACTCCAGATTAT 2595
781 GluLeuLeuSerAspProThrProGlyIleProSerProArg---AspGluLeuCysVal 799
2596 GAACCTGTCAGAAA---AATGACATGCTTACCACACATGAGATAGTGGTCCCTAGGATTT 2652
800 AlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLeuLys 819
2653 TCTGGTGCT-----TTTGAAGACAGGAGCCCTACACAGTTTGAAGAGAGACACTTGAAA 2706
820 TyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArgTyrAspPro 839
2707 TTTCTACAGCAACTTGGCAAGGGTAACTTTGGGAGTGTGGAGATGTCGCCCTATGACCCG 2766
840 LeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGlnHisSerValProAsp 859
2767 CTCAGGACAACTGGAGAGGTGATGCTTGAAGAGTCCAGCACACACGGAAGAG 2826
860 GlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHisSerAspPheIle 879
2827 CATCTCCGACACTTTGAAGGGAAATTTGAAATCTGAAATCTCTGACAGCATGACAACATC 2886
880 ValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArgLeuValMetGlu 899
```



Db 2887 GTGAAGTACAGGGAGTGTCTACAGTGTCTGGTGGCGGCAACCTTAAGATTAATTATGGA 2946  
 Qy 900 TyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArg---GlyLeuHisThr 918  
 Db 2947 TATTTGGCGTATGGAAGTTTACGAGACTATCTCCAAACAACATAAAGACGATAGATCAC 3006  
 Qy 919 AspArgLeuLeuLeuPheAlaTTrpGlnIleCysLysGlyMetGluTyrLeuGlyAlaArg 938  
 Db 3007 AAAAAAAGCTTCTGCAGTATACATCCAGATATGCAAGGCGATGAGTATCTTGGTACAAA 3066  
 Qy 939 ArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHisVal 958  
 Db 3067 AGGTATATCCACCGGATCTGGCAACAAGGAATATTTGGTAGAAATGGAACAGAGTT 3126  
 Qy 959 LysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAspTyrTyrValVal 978  
 Db 3127 AAAATAGGAGATTTTGGATTAAACCAAGTCTTGGCACAGGACAAAGATATCTACAAGTA 3186  
 Qy 979 ArgGluProGlyGlnSerProIlePheTyrTrpAlaProGluSerLeuSerAspAsnIle 998  
 Db 3187 AAGGAACACAGGTGAAGATCCCATATTTCTGGTATGCTCCGGAATCTTGCACAGAGCAAG 3246  
 Qy 999 PheSerArgGlnSerAspValTTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
 Db 3247 TTTTCTGTGGCTCAGATGTGTGAGCTTTGGTGGTCTTGTACGAACTTTTCCAGTAT 3306  
 Qy 1019 CysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGlyProGluArgGlu 1038  
 Db 3307 ATCAGAAAGAGCAAAAGCCACCTGTGGAATTTATGCGAATGATCGCAATGATAACAA 3366  
 Qy 1039 GlyProProLeuCys---ArgLeuLeuGluLeuLeuAlaGluGlyArgArgLeuProPro 1057  
 Db 3367 GGGCAGATGATCGTATTCCATTTTGATAGACTACTGAAGAACAACAGGAAGACTCCGAGA 3426  
 Qy 1058 ProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTTrpAlaProGluPro 1077  
 Db 3427 CCAGAGGTGCCCAGACGAGATTATGTGATCATGACAGAACTCTGGAAACAACATGTC 3486  
 Qy 1078 HisArgProAlaPheAlaThrLeuSerProGlnLeuAspProLeu 1093  
 Db 3487 AACCAACGTCCCTTTTCAGGGACCTTCCCTTCGGGTGGATCAATC 3534  
 RESULT 4  
 US-10-305-720-297  
 ; Sequence 297, Application US/10305720  
 ; Publication No. US20040010136A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.  
 ; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression  
 ; FILE REFERENCE: PA-0002-1 CON  
 ; CURRENT APPLICATION NUMBER: US/10/305,720  
 ; CURRENT FILING DATE: 2002-11-26  
 ; PRIOR APPLICATION NUMBER: 09/016,434  
 ; PRIOR FILING DATE: 1998-01-30  
 ; NUMBER OF SEQ ID NOS: 1490  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 297  
 ; LENGTH: 4078  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No. US20040010136A1 179527  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1) ... (4078)  
 ; OTHER INFORMATION: a, t, c, g, or other  
 US-10-305-720-297  
 Alignment Scores:  
 Pred. No.: 3,83e-257 Length: 4078  
 Score: 2594.00 Matches: 531  
 Percent Similarity: 65.94% Conservative: 193

Best Local Similarity:	48.36%	Mismatches:	343
Query Match:	44.27%	Indels:	32
DB:	12	Gaps:	14
US-09-397-967A-16 (1-1099) x US-10-305-720-297 (1-4078)			
QY	20	SerSerGluAlaGlyAlaLeuHisValLeuLeuPro-ProArgGlyProGlyProProGly	39
DB	447	TCAGTTCCTTCAGGT-----GTATCTTTACCATTCCTTGGGAAATCTGAGGCAGA	497
QY	39	nArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysValArgAlaAlaLys	59
DB	498	TTATCTGACCTTTCCATCTGGGGAGATGTTGGAGAAGAAATCTGTATTGCTGCTTCTAA	557
QY	59	sAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThrGluAspPhe	79
DB	558	AGCTTGTGGTATCACACCTGTATCATAAATATGTTTGTATTAGTGTGAACAGAAAG	617
QY	79	rCysTrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuVa	999
DB	618	GATCTGGTATCCACCAACCATGCTTCATATAGTAGTCAACAGGCATAATGTACT	677
QY	99	lTyrArgLeuArgPheTyrPheProAspTrpPhe-----GlyLeuGluThrCysHisnR	117
DB	678	CTACAGATAGATTTTACTTTTCCTCG-TGGTATTGCAGTGCAGCAACAGAGCCTATCG	736
QY	117	gPheGlyLeuArgLysAspLeuThrSerAlaIleLeuAspLeuHisValLeuGluHisLe	137
DB	737	GCATGGAATATCTCGAGGTCTGAAGCTCCTCTTCTTGATGACTTTGTTCATGCTTACCT	796
QY	137	uPheAlaGlnHisArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetIy	157
DB	797	CTTTGCTCAGTGGCGGCATGATTTTGTGCATGGATGGATAAAAGTACCTGTGACTCATGA	856
QY	157	sGluGlnGlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGl	177
DB	857	AACACAGAAAGATGCTCTTGGATGCAGCTGTAGATATGATGAGAAATAGCCAAAGAAA	916
QY	177	nAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuProProSe	197
DB	917	CGATCAAAACCCACTGGCCATCTATACTCTATCAGCTACAGACATCTTACCACAATG	976
QY	197	rLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgIleArgThrVa	217
DB	977	TATTTCGACAAAGATCCAAAGACTATCATATTTCACAGGAAGCAATTAAGTACAGATT	1036
QY	217	lValLeuAlaLeuLeuProCysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLy	237
DB	1037	TCGCAGATTATTTCAGCAATTCAGCCAAATGCAAGCCACTGCCAGAACTTTGAAACTTAA	1096
QY	237	sTyrIleLeuAspLeuGluArgLeuHisProAlaAlaThrThrGluThrPheArgValGl	257
DB	1097	GTATCTTATAAATCTGGAATCTGCACTCTGCAGTCTGCGCTTCTACACAGAAATTTGAAGTAA	1156
QY	257	yLeuProGlyAla-----GlnGluGluProGlyLeuLeuArgValAlaGlyAs	273
DB	1157	AGAACCTCGAAGTGGTCTCTCAGGTGAGGAGATTTTTCGAACCATTTATAACTCGAAA	1216
QY	273	pAsnGlyIleProTrpSerSer-----AsnAspGluLe	284
DB	1217	CGGTGGAAATTCAGTGGTCAAGAGGGAAACATAAAGAAAGTGAGACACTGACAGAACAGGA	1276
QY	284	uPheGlnThrPheCysAppPheProGluIleValAspValSerIleAsnGlnAlaProAr	304
DB	1277	TTTACAGTTATATGCGATTTTCTTATATATTATGATGTCAGTATTAAAGACAGCAACCA	1336
QY	304	gValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLe	324
DB	1337	AGAGGGT---TCAAATGAAAGCCAGTGTGAATCTATCCATAAGCAAGATGTTAAATATCT	1393
QY	324	uGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTy	344
DB	1394	GGAAATGGAATCTAGCTCATTAAGGGAGGCTTTGCTTCGTGCTCAATTAATGATGATA	1453

Alignment Scores:	
Pred. No.:	3,83e-257
Score:	2594.00
Length:	4078
Matches:	531
Percent Similarity:	65.94%
Conservative:	193







Db 2087 ATATGAACCAATGGTGTTCACAAAATCAGAAATGAGATTGATATTTAATGAAGCC 2146  
Qy 522 euGlyHisGlySerPheThrIysLlePheArgGlySerArgGluValValAlaSp---G 541  
Db 2147 TTGGCCAGGCACTTTTACAAAGATTTTAAAGGGCTACGAAGAGAGTAGGAGACTAGC 2206  
Qy 541 lYgluThrHisAspSerGluValLeuLeuValMetAspSerArgHisArg-AsnCys 560  
Db 2207 GTCAATGTCATGAACAGAGTTCCTTTAAAGGTTCTGGATTAAGCAGACAGAAACTAT 2266  
Qy 561 MetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHisLeuVal 580  
Db 2267 TCAGAGTCTTTCTTGAAGCATCAAGTATGATGAGCAACCTTCTCACAAGCATTTGGTT 2326  
Qy 581 LeuLeuHisGlyValCysMetAlaGlyAsp---SerIleMetValGlnGluPheValTyr 599  
Db 2327 TTAATAATGAGTAGTGTCTGTGAGACGGAATATTTCTGGTTCAGGAGTTTGTAATA 2386  
Qy 600 LeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSerTrpLys 619  
Db 2387 TTGGATCACTAGATACATATCTGAAAGAAATTAATAATTTGATATAATATATGAA 2446  
Qy 620 LeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGlyLeuPro 639  
Db 2447 CTTGAAGTTGCTTAAACAGTTGGCAGCGGCATGCAATTTCTAGAGAAACACCCCTATT 2506  
Qy 640 HisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGly-----AspGly 657  
Db 2507 CATGGCAATGTATGTGCAACACATATTTCTCTTATCAGAGAGAGAGACAGGACAGGA 2566  
Qy 658 AsnProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeuGlu 677  
Db 2567 AATCCCTCTTTCATCAAACTTAGTATGCTGGCATTTAGTATTAACGTTTGGCAAGGAC 2626  
Qy 678 MetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThrLeu 697  
Db 2627 ATTCTTCAGAGAGATACCATGGGTACCACTGATGATGATTAATAATCTTAAATTTA 2686  
Qy 698 CysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArgGly 717  
Db 2687 AATTGGCAACAGACAAATGAGTTTGGTACCACCTTTGTGGAAATCTGCAGTGGAGGA 2746  
Qy 718 Pro-AlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGlnGly 737  
Db 2747 GATAATACCTCTAGTGTCTCGGATCCCAAGAAAGCTACAATTTATGAGATAGGCA 2806  
Qy 737 yGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTy 757  
Db 2807 TCAGCTTCCTGCACCAAGTGGGCGAGAAATTACCACACCTTATAAATAATTGATGATTA 2866  
Qy 757 rAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIleTh 777  
Db 2867 TGAACCAAGATTTACAGGCTCTTTCAGAGCCATCATACGAGATCTTAACAGTTTGTTC 2926  
Qy 777 rSerAspTyrGluLeuLeuSerAspProThrProGlyLleProSerProArg---AspGl 796  
Db 2927 TCCAGATTATGAATTAACAGAA---AATGACATGTTACCAATATGAGATAGGTGC 2983  
Qy 796 uLeuCysValAlaAlaGlyAlaGlnLeuTyrAlaCysGln-AspProAlaIlePheGluGlu 816  
Db 2984 CCTAGGGTTTCTGGTACC-----TTTGAAGACCGGGGATCTCTACACAGTTTGAAGA 3037  
Qy 816 rgHisLeuLysTyrIleSerLeuLeuLysGlyAsnPheGlySerValGluLeuCys- 835  
Db 3038 GACATTTGAAATTTCTACAGCAACTTGGCAAGGGTAATTTTGGGAGTGTGGAGATGTGCC 3097  
Qy 836 ArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuHis 855  
Db 3098 CGGTATGACCTCTACAGACAACTGCTGGGAGGTGCTGCTGTAAAGAAAGCTTCAGAT 3157  
Qy 856 SerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHis 875  
Db 3158 AGTACTGAAGACGACCTTAAGACACTTGAAGGGAAATTGAATCTCGAAATCCCTACAG 3217

Qy 876 SerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArg 895  
Db 3218 CATGACAACTTGTAAAGTACAAAGGAGTGTCTACAGTGTGTGGTGGCGGTAATCTTAAA 3277  
Qy 896 LeuValMetGluTyrLeuPro-SerGlyCysLeuArgAspLeuLeuGlnArgHisArg-- 914  
Db 3278 TTAATATGGAATATTTACCATATATGGGAAGTTTACGAGACTATCTTCAAAAACATAAGA 3337  
Qy 915 -GlyLeuHisThrAspArgLeuLeuPheAla-TripGlnIleCysLysGlyMetGluT 934  
Db 3338 ACGGATAGATCATATAAACTTCTGCAGTACACATCTCCAGATATGCAAGGGTATGGAGT 3397  
Qy 934 yTrLeuGlyAlaArgCysValHisArgAspLeu-AlaAlaArgAsnIleLeuValGlu 953  
Db 3398 ATCTTGGTACAAAAGGTATATCCACAGGATCTGGCAACGAGAAATATATTGGTGGAG 3457  
Qy 954 SerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLys 973  
Db 3458 AACGAGAACAGAGATTAAAAATTGGGATTTTGGGTTAACCAAGTCTTGGCACAGACAA 3517  
Qy 974 AspTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTrpAlaProGluSer 993  
Db 3518 GAATACTATAAGTAAAGAACCTGTGTGAAGTCCCATATTTCTGGTATGCTCCAGAAATCA 3577  
Qy 994 LeuSer-AspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTy 1013  
Db 3578 CTGAGCNGAGAGCAAGTTTCTGTGGCCTCAGATGTTTGGAGCTTTGGAGTGTCTGTGA 3637  
Qy 1013 rGluLeuPheThrTyrCysAsp-LysSerCysSerProSerAlaGluPheLeuArgMet- 1032  
Db 3638 TGAACCTTTTACATACATTTGAGAAAAGTAAAGTCCACAGCGGAATTTATCGGTATGA 3697  
Qy 1033 MetGlyProGluArgGluGlyProProLeuCys---ArgLeuLeuGluLeuAlaGlu 1051  
Db 3698 TTGGGCAATGACAAACAGGACAGACAGATGCTGCTTCCATTGTAGAACATTTTGAAGAAT 3757  
Qy 1052 GlyArgGluLeuProProProThrCysProThrGluValGlnGlu-LeuMetGlnLe 1071  
Db 3758 AATGGAAGATTACCAAGACCAGATGGATGCCAGATGAGATGATGATGATGATGATGATG 3817  
Qy 1071 uCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeuSerProGlnLeuAs 1091  
Db 3818 ATGCTGAAACATAATGTAATCAACGCCCTCTTTAGGATCTAGCTCTTCGAGTGA 3877  
Qy 1091 pProLeu 1093  
Db 3878 TCAATA 3884

RESULT 6  
US-10-295-027-697  
; Sequence 697, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; PRIOR FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13



Qy 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla 502  
Db 1941 GTGGCTACTAAG-----AAAGCCAGAGGTGGAGCC--GTTACCCC 1982  
Qy 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeu 522  
Db 1983 ATGAGCCAGCTGAGTTTCGATCGATCTCTCAAGAGGATCTGGTCAGGGGAGCAGCCTT 2042  
Qy 523 GlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAspGlyGlu 542  
Db 2043 GGGAGAGCAGCAGAACACACATCTATTCTGGGACCTGTATGATTAACAGGATGACGAA 2102  
Qy 543 ThrHisaspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg 558  
Db 2103 GGAATCTTCTGAAGAGAGAGAGATAAAGTGTATCTCAAGTCTTAGACCCAGCCACAGG 2162  
Qy 559 AsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTrpProHis 578  
Db 2163 GATATTTCCCTGGCCTTCTTCGAGGAGCCAGCATGATGAGACAGGTCTCCACAAACAC 2222  
Qy 579 LeuValLeuLeuHisGlyValCysMetAlaGly----AspSerIleMetValGlnGluPhe 597  
Db 2223 ATCGGTACTCTATGGGTCTGTGTCTGGCAGCGTGGAGATATCATGTGGAGAGTTT 2282  
Qy 598 ValTyrLeuGlyAlaIleaspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer 617  
Db 2283 GTGGAAGGGGTCCTGTGATCTCTTCATGTCACCGGAAAGCGATGCTTACCACACCA 2342  
Qy 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly 637  
Db 2343 TGGAAATTCAAAGTTGCCAAACAGCTGGCCAGTGCCTGAGCTACTTGGAGGATAAAGAC 2402  
Qy 638 LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyClyAspGly 657  
Db 2403 CTGTGCTCATGGAATGTGTACTAAACCTCTCTGGCCCGTGGAGGCATCGACAGT 2462  
Qy 658 Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676  
Db 2463 GAGTGTGGCCCATTCATCAAGCTCAGTACCCCGGATCCCATACGCTGTGCTAGG 2522  
Qy 677 GluMetLeuThrAspArgIleProTrpValAlaProGluCysLeuGlnGluAlaGlnThr 696  
Db 2523 CAAGATGATGATGAACGATCCATGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2582  
Qy 697 LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArg 716  
Db 2583 CTGAGTGTGGCTGCTGACAAAGTGGAGCTTTGGAAACCCAGCTCTGGGAAATCTGCTACAAT 2642  
Qy 717 GlyProAlaHisIleThrSerLeuGluProAlaLysLysLysPheTyrGluAspGln 736  
Db 2643 GCGAGATCCCTTTGAAAGACAGACGCTGATGGAAGAGAGATCTTATGAAAGCCGG 2702  
Qy 737 GlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAla 756  
Db 2703 TGCAGCCAGTGCACCATCATGATTAAGAGCTGGCTGACCTCATGATGACCCGCTGATGAAC 2762  
Qy 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776  
Db 2763 TATGACCCCAATCAGAGGCTTTCTTCCGAGCCATCATGAGACATTAATAAGCTTGA 2822  
Qy 777 ThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGlu 796  
Db 2823 GACAGATCCAGATATTGTTTCAGAAAAAACCCAGCACTGAA----- 2867  
Qy 797 LeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArg 816  
Db 2868 -----GTGACCCCAACACATTTTGAAGAGCGC 2894  
Qy 817 HisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArg 836  
Db 2895 TTCTTAAGAGGATCCGTGATCTTGGAGAGGGGCGACCTTTGGGAGGTGAGTCTGCGAG 2954

Qy 837 TyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln---His 855  
Db 2955 TATGACCCAGAGGGGACAATACAGGGGAGCAGGTGCTGTAAATCTCTCAAGCCTGAG 3014  
Qy 856 SerValProAspGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHis 875  
Db 3015 AGTGGAGTAAACCATAGCTGATCTGAAAAGGAATTCGAGATCTTAAGAACTCTAT 3074  
Qy 876 SerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArg 895  
Db 3075 CATGAGAACATTGTGAGTACAAAGGAATCTGCACAGAACCGAGGAATGGTATTAAAG 3134  
Qy 896 LeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGlnArgHisArgGly 915  
Db 3135 CTCTATCGGAATTTCTGCTTCGGGAAGCTTAAAGGAATATCTTCCAAAGAAATAAGAAC 3194  
Qy 916 ---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMetGluTyr 934  
Db 3195 AAAATAAACCTTCAACAGCAGCTAAATATATGCCGTTCAGATTTGTAAAGGGATGGACTAT 3254  
Qy 935 LeuGlyAlaAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSer 954  
Db 3255 TTGGTCTCGCAATACGTTTACCGGAGCTTGGCAGCAAGAAATGTCTTGTGAGAGT 3314  
Qy 955 GluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAsp 974  
Db 3315 GAACACCAAGTGAATAATTGGAGACTTCGGTTTAAACCAAGCAATTTGAACCGATAGGAG 3374  
Qy 975 TyrTyrValValArgGluProGlyGlnSerProIlePheTyrTrpAlaProGluSerLeu 994  
Db 3375 TATTACCGTCAGAGATGACCGGACAGCCTGTGTTTGGTATGCTCCAGATGTTTA 3434  
Qy 995 SerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeuTyrGlu 1014  
Db 3435 ATGCAATCTAAATTTTATATTGCTCTGACGCTGCTCTTTTGGAGTCACTCTGCATGAG 3494  
Qy 1015 LeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034  
Db 3495 CTGTGACTTACTGTGATTCAGATTCTAGTCCCATGGCTTTGTTCTGAAAATGATAGC 3554  
Qy 1035 ProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGlyArgArg 1054  
Db 3555 CCAACCCATGGCAGCATGACATGACAGCTTGTGAATACCTTAAAGAGAGAAACGC 3614  
Qy 1055 LeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCysTrpAla 1074  
Db 3615 CTGCGGTGCGCCACCTAACTGTCCAGATGAGTTTATCAACTTATGAGGAATGCTGGAA 3674  
Qy 1075 ProGluProHisAspArgProAlaPheAlaThrLeu 1086  
Db 3675 TTCCAACCATCCAATCGGACAAAGCTTTTCAGAACCTT 3710

## RESULT 7

US-10-084-817-230  
; Sequence 230, Application US/10084817  
; Publication No. US20030119009A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Stuart  
; APPLICANT: Ted G. Nuchtern  
; APPLICANT: Sharon E. Plon  
; APPLICANT: Jason M. Shohet  
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION  
; FILE REFERENCE: PA-0046 US  
; CURRENT APPLICATION NUMBER: 60/270,784  
; PRIOR FILING DATE: 2002-02-25  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 365  
; SOFTWARE: PERL Program  
; SEQ ID NO 230  
; LENGTH: 5073  
; TYPE: DNA  
; ORGANISM: Homo sapiens

```

; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 903876.1
US-10-084-817-230

```

Alignment Scores:

Pred. No.:	4,51e-188	Length:	5073
Score:	1927.00	Matches:	443
Percent Similarity:	53.92%	Conservative:	189
Best Local Similarity:	37.80%	Mismatches:	384
Query Match:	32.88%	Indels:	156
DB:	15	Gaps:	22

US-09-397-967A-16 (1-1099) x US-10-084-817-230 (1-5073)

QY	17	SerLeuSerSerGluAlaGlyAlaLeuHisValLeu---	LeuProProArgGlyPro	35
Db	354	ACCTGGAGGCCCTGACCCAGGGGTGAAGTATCTTCATCTGTGGACGGGACGCC	413	
QY	36	GlyProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysVal	55	
Db	414	-----CTCCGGCTGGCAGGTGACAGTACACAGCAGAGNACTGTGCATC	458	
QY	56	ArgAlaAlaLysAlaCysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThr	75	
Db	459	AGGGCTGCACAGGATGCCGATCTCTCTCTTTGTCAACCTTTTGCCTGTATGAC	518	
QY	76	GluAspPheSerCysTrpPheProSerHisIlePheCysIleGluAspValAspThr	95	
Db	519	GAGAACCAAGCTCGTATGCTCCAAATCGCACCATCACCGTTGATGACAGATGCTC	578	
QY	96	GlnValLeuValTyrArgLeuArgPheTyrPheProAspTrpPheGlyLeu---	112	
Db	579	CTCCGGCTCCACTACCGGATGAGGTTCTATTACCAATTTGGCATGGAACCAACGACAT	638	
QY	113	---GluThrCysHisArgPheGlyLeuArgLysAspLeu---	124	
Db	639	GAGCAGTCAGTGTGGCGTCATTCTTCAAAGAACAGACAAAAATGGCTACGAGAAAAAAG	698	
QY	125	-----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis	141	
Db	699	ATTCCAGATGCAACCCCTCTCTTGATGCCAGCTCATCTGGATATCTGTTGCTCAGGA	758	
QY	142	ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln---	159	
Db	759	CAGTATGATTGTGTGAATGCCTGGCTCCTATTTCGAGACCCCAAGACCGAGCAGATGGA	818	
QY	160	-----GlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArg	175	
Db	819	CATGATATTGAAACGAGTGTCTAGGATGGCTGTCTCTGGCCATCTCACACTATGCCATG	878	
QY	176	GluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuPro	195	
Db	879	ATGAAGAAGATGCGATGGCCAGAACTGCCACAGGACATCAGTACACGGCATATATTC	938	
QY	196	ProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgIle---	213	
Db	939	GAAACATTGAATAAGTCCATCAGACAGAGAACCTTCTCACCGAGTCCGATAAATAAT	998	
QY	214	-----ArgArgThrValValLeuAlaLeuLeuPro	223	
Db	999	GTTTTCAAGTATTCCTAAAGGAATTTACACAGACCATTTGTGACAGCAGCGGTG---	1055	
QY	224	CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu	243	
Db	1056	-----TCCACGATGACCTGAAGTCAATATCTTCGCTACCTTGGTAA	1097	
QY	244	ArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu	263	
Db	1098	ACTTTGACAAAAATTACGCTGCTGAAATTTT-----	1130	
QY	264	GluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyr---SerSerAsnAsp	282	

1131	Db	GAGACTTCATGTTACTGATTTCATCAGAAAATGAGATGAATTGGTTTCATTTCGAATGAC	1190
282	QY	-----	282
1191	Db	GGTGGAAACGTTCTCTACTACGAAGTGATGGTGACTGGGAATCTTGGAAATCCAGTGGAGG	1250
282	QY	-----	282
1251	Db	CATAAACCAAAATGTTGTTTCTGTTGAAAAGGAAAAATAAACTGAAGCGGAAAAAACTG	1310
283	QY	-----	283
1311	Db	GAAAATAAACACAAGAAGATGAGGAGAAAAACAAGATCCGGGAAGAGTGGCAACAATTTT	1370
289	QY	CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAla	308
1371	Db	TCTTACTTCCCTGAAATCACTCACTTGTATATAAGAGTCT	1412
309	QY	GlyGluHisArgLeuValThrValArgMetAspGlyHisIleLeuGluAlaGluPhe	328
1413	Db	-----GTGGTCAGCATTAACAAGCAGGACACAAGAATAATGGAATCAAGCTC	1460
329	QY	ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle	348
1461	Db	TCTTCCCAAGAGGAGGCTTGTCTTTGTCTCCCTGGTAGATGGCTACTTCCGGCTCACA	1520
349	QY	CysAspSerArgHisIstYrPheCysLysGluValAlaProArgLeuLeuGluGlu	368
1521	Db	GCAGATGCCCATCACTTCGACCGAGTGGCCCCCGTTGATCGTCCACAACATA	1580
369	QY	AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAla	388
1581	Db	CAGAAATGGCTGTCATGGTCCAAATCTGACAGAATACCCCATCAATAATTTGGCGCAAGAA	1640
389	QY	GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu	408
1641	Db	GGAAGCGAGGAGGGAGTACGTCTCGGTGAGGTGCACCGCACTTGCACACATCTCTC	1700
409	QY	LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys	422
1701	Db	ATGACCGTCACCTGCTTTGAGAAGTCTGACGAGTGCAGGGTGCACGAAGACGAGTTC	1760
423	QY	GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro	442
1761	Db	AACTTTTCAGATC---GAGGTGCAGAAGGGCGGCTACAGTCTGCACGGTTTCGACCGCAGC	1817
443	QY	HisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGly	462
1818	Db	TTCCCCAGCTTGGGAGCACTCATGAGCCACCTCAAGAACGACATCTCTGCGCAGGATAAC	1877
463	QY	AlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIle	482
1878	Db	ATCAGCTTCATGCTAAAACGCTGCTGCCAGCCCAAGCCCCGAGAATCTTCCAACCTGTGT	1937
483	QY	ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla	502
1938	Db	GTGGCTACTAAG-----AAAGCCCAGGAGTGCACGCC---GTCTACCC	1979
503	QY	LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeu	522
1980	Db	ATAGCCAGCTGAGTTTCGATCGGATCTCTCAAGAAGGATCTGGTGCAGGGCGAGCACCTT	2039
523	QY	GlyHisGlySerPheThrLysIlePheArgGlySerArgGluValValAspGlyGlu	542
2040	Db	GGGAGAGGCACGAGACACACATCTATTCTTGGGACCCTGATGGATTACAAGATGACGAA	2099
543	QY	ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg	558
2100	Db	GGAACCTTCTGAAGAGAAGAAGATAAAGATGATCTCTCAAGTCTTAGACCCCGACCCACAGG	2159
559	QY	AsnCysMetGluSerPheLeuGluAlaSerLeuMetSerGlnValSerTyrProHis	578
2160	Db	GATATTTCCTGGCTTCTTCAGAGCGCAGCATGATGACACAGGCTCTCCCAACACAC	2219



QY 81 TrpPheProSerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyr 100  
DB 528 TGGTACGCTCCGAACGAATCATCACTGTGGATGACAAAACGCTCTCTCGGCTCCACTAC 587  
QY 101 ArgLeuArgPheTyrPheProAspTrpPheGlyLeu-----GluThrCysHis 116  
DB 588 CGCATGAGGTTCTACTTTACCAACTGGCAGCAACCAATGACAAACAGCTGTATG 647  
QY 117 ArgPheGlyLeuValAspLeu-----ThrSer 126  
DB 648 CGACATTCTCCAAAGACAGCAAAAACCGCTATGAGAAAGAGGTTCAGAAACCAACC 707  
QY 127 AlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuVal 146  
DB 708 CCATCTCCTGTATGTCAGTTCACTGGAGTATCTGTTGACAGGACAGTATGATTC 767  
QY 147 SerGlyArgLeuProValGlyLeuSerMetIysGluGln-----Gly 160  
DB 768 AAATGCTGGCTCCATTTCCGGACCCCAAGACGGAGCAAGACGGACATGATATTGAAAT 827  
QY 161 GluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArg 180  
DB 828 GAGTGCCTGGCATGGGGTCTCTGGCCATCTCCCACTATGCCATGATGATGAAAGATGCAG 887  
QY 181 ProGlyGluLeuLeuLysThrValSerTyrIysAlaCysLeuProProSerLeuArgAsp 200  
DB 888 TTGCCGGAACCTTCCCAAGACATCAGCTACAGCGATATATTCCAGAAACATTTGAATAA 947  
QY 201 ValIleGlnGlyGlnAsnPheValThrArgArgIle----- 213  
DB 948 TCATCAGACAGAGAACCTTCTTACAGGATGCGAATAATAATGTTTCAAGGATTTTC 1007  
QY 214 -----ArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeuPro 228  
DB 1008 TTGAAGGAATTTAACAACAGACCATC-----TGTGAC 1040  
QY 229 GlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu-----ArgLeuHis 246  
DB 1041 AGCAGTGTGCATGACCTGAAGGTGAATACCTGCTACCTTGGAACTTCTACATTGACA 1100  
QY 247 ProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGluProGly 266  
DB 1101 AACATTTAGGACTGAAATATTGTAGACTTCTATGCTACTGATTTTCATCAGAAATGAA 1160  
QY 267 LeuLeuArg-----ValAlaGly 272  
DB 1161 TTGAGTCGATGCTTCCGATGACATGCGCAATGTTCTCTATGAGTCTGGTGGATGGA 1220  
QY 273 AspAsnGlyIleProTrpSerSerAsnAsp----- 282  
DB 1221 AATCTCGGATCCAGTGGCGGCAACCAAAATGTTGTTCTCTGTTGAAAGGAAATAAT 1280  
QY 282 ----- 282  
DB 1281 AAACCTAAGCGGAAAACCTGGAATATAATAACACAAAGAGGATGATGAGAAACAAA 1340  
QY 283 -----GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsn 300  
DB 1341 CTCGGGAAGAGTGAACAAATTTTCTATTTCCCTGAAATACCCACATTTGTAATAAG 1400  
QY 301 GlnAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAsp 320  
DB 1401 GAGTCT-----GTGTCAGCATTAACAACAGGAC 1430  
QY 321 GlyHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeu 340  
DB 1431 AACAAAAACATGGAATCTCAAGCTCTCTCTCGAGAGGAAGCCTTGTCTTGTGCTCCG 1490  
QY 341 ValAspGlyTyrPheArgLeuIleCysAspSerArgHisTyrPheCysIysGluValAla 360  
DB 1491 GTGGATGGTACTTCGGGTCTACTGCAGATGCCACCATTTACTCTGTACTGTGGCT 1550  
QY 361 ProProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPhe 380

DB 1551 CCCCACCTGATGTCCACAATATACAGACGGCTGCCAGGTCCTCAATCTGCACAAATAT 1610  
QY 381 AlaIleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSer 400  
DB 1611 GCCATCAATAAGTCGGCAGGAAGGAGTGAAGAGGGATGTACGTCTGAGGTGGAGC 1670  
QY 401 ProGlnAspTyrAspSerPheLeuLeuThr-----AlaCysValGlnThrPro-----Leu 417  
DB 1671 TGCACCGACTTTGACACATTTCTATGACTGTCACTGCTTGAAGAGTCTGAGGTATTG 1730  
QY 418 Gly-----ProAspTyrLysCysLeuLeuArgGlnAspProSerGlyAlaPheSer 435  
DB 1731 GGTGGCCAGAAGCAGTTCAAGAACTTTTCAGATT---GAGGTACAGAAGGGCGCTACAGC 1787  
QY 436 LeuValGlyLeuSerGlnProHisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsn 455  
DB 1788 CTGCATGGCTCTATGGACCACTTTCCAGCCTTCGAGACCTCATGAACCACTCTCAGAG 1847  
QY 456 SerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProArgPro 475  
DB 1848 CAGATCTCTGCGCACGACACATAAGCTTTGTGCTGMAACGATGCTGTCAAGCTTAAAGCT 1907  
QY 476 LysGluLysSerAsnLeuIleValValArgArgGlyCysAsnProAlaProAlaProGly 495  
DB 1908 CGAGAAATCTCAATCTGCTGTCAGTACCACTAAG-----AAAGCCCAAGAG 1952  
QY 496 CysSerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAspSer 515  
DB 1953 TGGCAGCCT---GTCATCTCATGAGCAGCTGAGCTTTGATCGGATCTTTAAGAAAGAT 2009  
QY 516 LeuGluTrpHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySerArg 535  
DB 2010 ATTATACAGGTGAGCACCTTTGCAGAGGCACAAACACATATCTATTCTGGGACCTG 2069  
QY 536 ArgGluValValAsp-----GlyGluThrHisAspSerGluValLeuLeuLys 551  
DB 2070 CTGAGTACAGAGTGAAGAGAAATGCTGAAGAGAGAGAAAGATGATCTCTCAAA 2129  
QY 552 ValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMet 571  
DB 2130 GTCTAGACCCAGCCAGCCGAGCATCTCTCGGCTTCTTTGAGGCTGTACATGATG 2189  
QY 572 SerGlnValSerTyrProHisLeuValLeuHisGlyValCysMetAlaGly---Asp 590  
DB 2190 AGACAGGTTTCCACAAACATATAGTGTACTCTACGGCTGTGTGTCGAGATGTGAA 2249  
QY 591 SerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArg 610  
DB 2250 AATATCATGTGGAAGAGTTTGTGGAGGGGGCGGCTTGGATCTCTTCATGCACCGGAAA 2309  
QY 611 GlyHisLeuValSerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeu 630  
DB 2310 AGTGATGCGCTTACTACCCCTGGAAAGTTCAAGTTGCCAAACAGCTGGCCAGTGGCCTG 2369  
QY 631 AsnTyrLeuGluAspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeu 650  
DB 2370 AGTTACTTGAAGATAAAGACCTGTTTCATGGAATGTGTGCACATAAAACCTCTCTG 2429  
QY 651 AlaArgGluGlyGlyAspGlyAsn---ProProPheIleLysLeuSerAspProGlyVal 669  
DB 2430 GCCCGTAGGGATTTGACAGTGCATTTGCCCGGCTTCATCAAGCTTAGTACCGCTGCATC 2489  
QY 670 SerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTrpValAlaProGlu 689  
DB 2490 CCAGTCTCTGTCTCACCGCAAGAGTGCATAGAGCGAATCCCTGAGTCTCCTGAG 2549  
QY 690 CysLeuGlnGluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThr 709  
DB 2550 TGTGTTGAAGACTCCAAAGAACCTGAGTGTGGTGTGTGACAGTGGAGCTTTGGAAACAGC 2609  
QY 710 ThrTrpGluValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLys 729



Db 2610 CTCTGGGAATCTCTCTACACGGAGAGATTCTCTCAAGACAGACCCCTCATTTGAGAA 2669

Qy 730 LeuLysPheTyrGluAspGlnGlyLeuProAlaLeuLysTrpThrGluLeuAlaGly 749

Db 2670 GAGAGGTTTATGAAGCGCTGCAGCGCTGTGACTCTCCATCTTGCAGAGGCTAGCTGAC 2729

Qy 750 LeuLeuThrGlnCysMetAlaTyrAspProGlyArgArgProSerPheArgAlaLeuLeu 769

Db 2730 CTCTAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2789

Qy 770 ArgAspLeuAsnGlyLeuLeuThrSerAspTyrGluLeuLeuLeuSerAspProThrProGly 789

Db 2790 AGGCACATTAAACCTGAGAGGAGAGATCCAGACATTGTTTCAGAA 2837

Qy 790 IleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAsp 809

Db 2838 --AAGCAGCCCAACACAGAGGTG-----GAC 2861

Qy 810 ProAlaIlePheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPhe 829

Db 2862 CCCACTCACTTTTGAAGCGTTCCTGAAGAGATTCTGCTGAGAGAGGTCACATT 2921

Qy 830 GlySerValGluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAla 849

Db 2922 GGGAGGTTGAGCTCTGAGATATGATCTGAGGGAGACACACAGGGGAGCAGGTAGCT 2981

Qy 850 ValLysGlnLeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluLe 868

Db 2982 GTCAAGTCCCTGAAGCTGAGAGTGGAGTGAACCATAGCTGCTGAAGAGAGATA 3041

Qy 869 GlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGly 888

Db 3042 GAGATCTTACCGAAGCTCTACCATGAGACATTGTGAAGTACAAAGGAATCTGCATGAA 3101

Qy 889 ProGlyArgGlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAsp 908

Db 3102 GACGGAGGAAGTGTATCAAGCTCATCATGAGATTCTGCTTCGGGAAGCCTAAAGGAG 3161

Qy 909 LeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTrpGln 927

Db 3162 TATCTGCCAAGAATAAAGAACAAATCAACCTCAACAGCAGCAGCTAAATATGCTCATCC 3221

Qy 928 IleCysLysGlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAla 947

Db 3222 ATTTGTAAAGGGATGACTACTTGGGTCTCGCAATACGTTTCCCGGAGCTTAGCAGCA 3281

Qy 948 ArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLys 967

Db 3282 AGAATGTCTTGTGTGAGTGGATGATCAAGTGAAGATCGAGACTTTGGTTTAAACCAA 3341

Qy 968 LeuLeuProLeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePhe 987

Db 3342 GCAATTGAACCCGATAGGAGTACTACAGTCAAGGACGACCGGGACAGCCAGTGTT 3401

Qy 988 TrpTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSer 1007

Db 3402 TGTACCTCCGGAATGTTAATCCAGTGAATTTTATATGCTCTGATGCTGTCT 3461

Qy 1008 PheGlyValValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAla 1027

Db 3462 TTTGGAGTGACCTGCAGAGCTGCTCATTACTGTGACTCAGATTTTGTCCCATGGCC 3521

Qy 1028 GluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGlu 1047

Db 3522 TTGTCTCTGAAATGATAGGCCCAACTCATGCGCAGATGACAGTGAACCGCTGTGAAG 3581

Qy 1048 LeuLeuAlaGluGlyArgArgLeuProProProProThrCysProThrGluValGlnGlu 1067

Db 3582 ACTCTGAAGAGGAAGGCTGTGCTCATGCTCCACCACTGCTCTGATGAGTTTATCAG 3641

Qy 1068 LeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086

Db 3642 CTTATGAGAAATGCTGGGAATTCCAACTCAACCGGACAACTTTTTCAGAACCTT 3698

RESULT 9

US-09-880-107-2379  
 ; Sequence 2379, Application US/09880107  
 ; Patent No. US20020142981A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Horne, Darci T.  
 ; APPLICANT: Vockley, Joseph G.  
 ; APPLICANT: Scherf, Uwe  
 ; APPLICANT: Gene Logic, Inc.  
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 ; FILE REFERENCE: 44921-5028-WO  
 ; CURRENT APPLICATION NUMBER: US/09/880,107  
 ; CURRENT FILING DATE: 2001-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/211,379  
 ; PRIOR FILING DATE: 2000-06-14  
 ; PRIOR APPLICATION NUMBER: US 60/237,054  
 ; PRIOR FILING DATE: 2000-10-02  
 ; NUMBER OF SEQ ID NOS: 3950  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2379  
 ; LENGTH: 3541  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M64174  
 US-09-880-107-2379

Alignment Scores:  
 Pred. No.: 3541  
 Score: 443  
 Percent Similarity: 53.66%  
 Best Local Similarity: 37.67%  
 Query Match: 32,49%  
 DB: 24  
 Gaps: 24

US-09-397-967A-16 (1-1099) x US-09-880-107-2379 (1-3541)

Qy 17 SerLeuSerSerSerGluAlaGlyAlaLeuHisValLeu---LeuProProArgGlyPro 35

Db 121 AACCTGGAGGCCCTGAGCCAGGGGTGGAAGTGTCTTCTATCTGTCGACAGAGGAGCCC 180

Qy 36 GlyProProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuVal 55

Db 181 -----CTCGGGTGGCAGTGGAGAGTACACAGCAGGAACTGTGCATC 225

Qy 56 ArgAlaAlaLysAlaCysGlyIleLeuProValTyrHisSerLeuPheAlaLeuAlaThr 75

Db 226 AGGGCTGCACAGCAGTCCGCTATCTCTCTCTTTGTCAACACTCTTTGCCCTGTATGAC 285

Qy 76 GluAspPheSerCysTrpPheProProSerHisIlePheCysIleGluAspValAspThr 95

Db 286 GAGAACACCAAGCTCTGGTATGCTCCAAATCGCACCATCACCGTTGATGACAAGATGCC 345

Qy 96 GlnValLeuValTyrArgLeuArgPheTyrPheProAspTrpPheGlyLeu----- 112

Db 346 CTCGGCTCCATACCGGATGAGGTCTATTTCACCAATTGGCATGGACCAACAGCAAT 405

Qy 113 ---GluThrCysHisArgPheGlyLeuArgLysAspLeu----- 124

Db 406 GAGCAGTCACTGTGGCTGCTATCTCCAAAGACAGCAAGAAATGGCTACGAGAAAAAAG 465

Qy 125 -----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141

Db 466 ATCCAGATCAACCCCTCTCTCTGATGTCGAGCTCACTGGAGTATCTGTTTGTCTCAGGA 525

Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----- 159

Db 526 CAGTATGATTGGTGAATGCTGCTGCTTATTCGAGACCCCAAGACCGACGAGATGGA 585

Qy 160 -----GlyGluPheLeuSerLeuAlaValLeuAspLeuGlnMetAlaArg 175

Db 586 CATGATATTGAGAACGAGTGTCTAGGATGGCTCTCTGCGCATCTCACATCTGCGCATG 645



QY	176	GluGlnAlaGlnArgProGlyGlyLeuLeuIysThrValSerTyrLysAlaCysLeuPro	195
DB	646	ATGAAGAAGATGCAGTTGCCAAGACTGCCAAGGACATCAGCTACAAGCGATATATTCCA	705
QY	196	ProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgArgIle----	213
DB	706	GAACAAITTAAGTCATCCAGACAGAGAACCTTCTCACCAAGATCGGATAAATAAT	765
QY	214	-----ArgArgThrValValLeuAlaLeuLeuPro	223
DB	766	GTTTTTCAAGGATTTCTTAAAGCAATTTAAACAAGACCATTGTGGACAGACGGTG--	822
QY	224	CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu	243
DB	823	-----TCCACGCATCAGCTCGAAGTGAATAACTTGGCTACCTTGGCATTTGAA	864
QY	244	ArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu	263
DB	865	ACTTTGACAAAACATTACGGTGCTGAATAATTT-----	897
QY	264	GluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrp--SerSerAsnAsp	282
DB	898	GAGACTTCCATGTTACTGATTTTCATCAGAAAAATGAGATGAATTGGTTTCATTCGAATGAC	957
QY	282	-----	282
DB	958	GGTGGAAACGTTCTCTACTCGAAGTAGTGATCTGGGAATCTTGGAAATCCAGTGGAGG	1017
QY	282	-----	282
DB	1018	CATAAACCAATGTTGTTCTGTTGAAAAGGAAAAAATAAATCAAGCGGAAAAAACTG	1077
QY	283	-----GluLeuPheGlnThrPhe	288
DB	1078	GAAATAAAGACAAGAAGGATGAGGAGAAAAACAAGATCCGGGAAGAGTGGAAACAATTTT	1137
QY	289	CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGilyProAla	308
DB	1138	TCATTCTTCCTGAAATCACTCATTGTATAAAGAGTCT-----	1179
QY	309	GlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhe	328
DB	1180	-----GTGGTCAGCATTAACAAGCAGGACACAGAAAAATCGAATCGAAGCTC	1227
QY	329	ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle	348
DB	1228	TCTTCCCACAGAGAGGCTTGCTCTCTGTGTCCCTGGTAGATGGCTACTTCCGGCTCACA	1287
QY	349	CysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu	368
DB	1288	GCAGATGCCCATCATTAACCTCTGCACCGAGTGGCCCCCGTGTGATCGTCCACAACATA	1347
QY	369	AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuLysAlaAla	388
DB	1348	CAGAATGGCTGTCATGTGTCCAATCTGTACAGAATACGCCATCAATAAATTCGGCAAGAA	1407
QY	389	GlySerLeuProGlyThrThrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu	408
DB	1408	GGAAGCAGAGGGGATGTACGTCTGAGTGGAGCTGCACCGCACTTTGACAACTCCTC	1467
QY	409	LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys	422
DB	1468	ATGACCGTCACCTGCTTTTGAGAGTCTGACGAGTGCAGGGTGCACAGACGAGTTCNAG	1527
QY	423	GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro	442
DB	1528	AACATTTTCAGATC---GAGGTGCAGAAGGGCGCTACAGTCTGCACGGTTCGGACCGCAGC	1584
QY	443	HisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGly	462
DB	1585	TTCCCCAGCTTTGGGAGACCTCATGAGCCACTCAAGAAGACAGATCTCTCGCAGCGGATAAC	1644

Qy	463	AlaLalaLeuTyrl	LeuThrsrCysCyseA	AlaProArgProlysGluIysSerAsnLeuile	489
Db	1645	ATCAGCTTCATGTTAAACGGCTGTCCAGCCCAAGCCCCGAGAAATCTCCAACCTCGT	1704	::: ::: :::	
Qy	483	VallValArgArgGlyCysAsnProAlaProAlaProAlaProArgCysSerProSerCysCyseAla	502	     	
Db	1705	GTGGCTACTAAG-----AAAGCCCAGGAGTGCCAGCCCC--GTCTACCCC	1746	     	
Qy	503	LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTyrHisGluAsnLeu	522	     	
Db	1747	ATGAGCCAGCTGAGTTTCGATCGCATCTCAAGAGGATCTGGTGCAGGCGGAGCACCTT	1806	     	
Qy	523	GlyHisGlySerPheThrLysIlePheArgGlySerArgGluValValAspGlyGlu	542	     	
Db	1807	GGGAGAGGACAGAGAACAACATCTATTCTGGGACCCTGTATGATTACAGGATGACGA	1866	     	
Qy	543	ThrHisaspserGlu-----ValLeuLeuIysValMetaspSerArgHisArg	558	     	
Db	1867	GGAACTTCTGGAAGAGAGAAGATAAAAAGTGATCTCTCAAAGTCTTTAGACCCACGCCACAGG	1926	     	
Qy	559	AsnCysMetGlnSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis	578	     	
Db	1927	GATATTCCCTGGCCCTTCTCGAGCCAGCAGATGATGAGACAGGTCTCCCAACAACAC	1986	     	
Qy	579	LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe	597	     	
Db	1987	ATCGTGTACCTTATGGCGTCTGTGCCGCGACGTGGAGAATATCATGTGTGAAGAGTTTT	2046	     	
Qy	598	ValtyrLeuGlyAlaIleAspMetTyrlLeuArgIysArgGlyHisLeuValSerAlaSer	617	     	
Db	2047	GTGGAAGGGGTCTCTGGATCTCTTCATGCACCGGAAAAAGTAGTGTCTTACCACCA	2106	     	
Qy	618	TripLysLeuGlnValThrLysGlnLeuAlatyAlaLeuAsnTyrlLeuGluAspLysGly	637	     	
Db	2107	TGGAAATTCAAAGTTGCCAACAGCTGGCCAGTGCCCTGAGCTACTTGGAGGATTAAGAC	2166	     	
Qy	638	LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGly	657	     	
Db	2167	CTGTCCATGGAATGTGTACTAMAAACCTCTCTCTGCCCGTGAGGGAATCGACAGT	2226	     	
Qy	658	Asn---ProPheIleIysLeuSerAspProGlyValSerProThrValLeuSerLeu	676	     	
Db	2227	GAGTGTGGCCATTCAACAAGCTCAGTGACCCCGGCATCCCAATACGGTGTCTTAGG	2286	     	
Qy	677	GluMetLeuThrAspArgIleProTripValAlaProGluCysLeuGlnGluAlaGlnThr	696	     	
Db	2287	CAAGAAATGCATTGACGAATCCCATGGATTCTCTGAGTGTGTGAGAGACTCCAAGAC	2346	     	
Qy	697	LeuCysLeuGluAlaAspLysTyrGlyPheGlyAlaThrThrTrpGluValPheGlnArg	716	     	
Db	2347	CTGAGTGTGGCTGTGACAAAGTGAGCTTTGGAACCCACCGCTCTGGGAAATCTGCTACAAT	2406	     	
Qy	717	GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuIysPheTyrgluaspGln	736	     	
Db	2407	GGCAGATCCCTTTGAAAGACAAGACGCTGATTGAGAAAGAGAGATCTCATGAAAGCCGG	2466	     	
Qy	737	GlyGlnLeuProAlaLeuLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAla	756	     	
Db	2467	TGCAGGCCAGTGACACCATCATGTAGGAGCTGGCTGACCTCATGACCCGCTGCATGAAC	2526	     	
Qy	757	TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeulle	776	     	
Db	2527	TATGACCCCAATCAGAGGCGCTTCTCCGAGCCATCATGAGAGACATTAAATAGCTTGA	2586	     	
Qy	777	ThrSerAspTyrgluLeuLeuSer-----AspProThrProGlylleProSer	792	     	
Db	2587	GACGAGAAATCCAGATATTGTTCCGAAAAAACCACGCCAACTGAA-----	2634	     	
Qy	793	ProArgspgluleucysValAlaGlyAlaGlnLeuTyrlAlaCysGlnAspProAlaIle	812	     	
Db	2635	-----GTGGACCCCAACAT	2649	     	
Qy	813	PheGluGluargHISleuLysTyrlleSerLeuLeuGlyLysGlyAsnPheGlySerVal	832	     	

```

|||||...|||||...|||||...|||||...|||||...|||||...
2650 TTTGAGAGCGCTCTCTAAAGAGGANTCGTGACCTGGAGAGGCGCATTGGGAGAGTT 2709
Qy
833 GluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValIysGln 852
Db
2710 GAGCTCTGCAGGTATGACCCC---GAAGACAATACAGGGAGCAGGTGGCTGTAAATCT 2766
Qy
853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnLeu 871
Db
2767 CTGAAGCTGAGAGTGGAGGTAAACCATAGCTGATCTGAAAGGAAGAAATCGATCTTA 2826
Qy
872 LysAlaLeuHisSerAspPheIleValIysTyrArgGlyValSerTyrGlyProGlyArg 891
Db
2827 AGGAACCTCTATCATGAGAACATTGTGAAGTACAAAGGAATCTGCACAGAAGCGGAGA 2886
Qy
892 GlnSerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGln 911
Db
2887 AATGGTATTAAAGCTCATCATGAGAAATTTCTGCTTCGGGAAGCCTTTAAGGAATATCTTCCA 2946
Qy
912 ArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTyrGlnIleCysLys 930
Db
2947 AAGATAAGACAAATAAATACCTCAACAGCAGCTAAATATGCGCTTCAGATTGTGAAG 3006
Qy
931 GlyMetGluTyrLeuGlyAlaArgCysValHisArgAspLeuAlaAlaAlaArgSerLeu 950
Db
3007 GGGATGGACTATTGGGTTCTCGCAATACGTTACCGGACTTGGCAGCAAGAAATGTC 3066
Qy
951 LeuValGluSerGluAlaHisValIysIleAlaAspPheGlyLeuAlaIysLeuLeuPro 970
Db
3067 CTTGTTGAGAGTGAACACCAAGTGAATTTGGAGACTTCGGTTTAAACCAAGCAATTGAA 3126
Qy
971 LeuGlyLysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTyrAla 990
Db
3127 ACCGATAAGGATATTACACCTCAAGGATGACCGGACGACCTGTGTTGGTATGCT 3186
Qy
991 ProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTyrSerPheGlyVal 1010
Db
3187 CCAGATGTTTAAATCAATCTAAATTTATATTATGCTCTGACGCTGGTCTTTGGAGTC 3246
Qy
1011 ValLeuTyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeu 1030
Db
3247 ACTCTGCATGAGCTGCTGACTTACTGTGATTGAGATTCAGATTCCTAGTCCATGGCTTTGTCCTG 3306
Qy
1031 ArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeuLeuAla 1050
Db
3307 AAAATGATAGGCCCAACCCATGGCCAGATGACAGTCACAGACTGTGATACGTTAATAA 3366
Qy
1051 GluGlyArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGln 1070
Db
3367 GAAGGAAACGCTGCGTGGCCACCTAACTGTCCAGATGAGGTTTATCAGCTTATGAGA 3426
Qy
1071 LeuCysTyrAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086
Db
3427 AAATGCTGGGAATTCCAACCATCCATCGACAAAGCTTTTCAAGAACCTT 3474

```

RESULT 10  
 US-09-873-367C-163  
 ; Sequence 163, Application US/09873367C  
 ; Publication No. US20030165839A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Young, Paul  
 ; APPLICANT: Soppet, Daniel  
 ; APPLICANT: Endress, Gregory  
 ; APPLICANT: Augustus, Meena  
 ; APPLICANT: Ebner, Reinhard  
 ; APPLICANT: Carter, Kenneth  
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
 ; FILE REFERENCE: 689290-84  
 ; CURRENT APPLICATION NUMBER: US/09/873,367C  
 ; CURRENT FILING DATE: 2003-04-29  
 ; PRIOR APPLICATION NUMBER: U.S. 60/236,891  
 ; PRIOR FILING DATE: 2000-09-29

```

; PRIOR APPLICATION NUMBER: U.S. 60/236,842
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: U.S. 60/244,867
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: U.S. 60/245,084
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 1067
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 163
; LENGTH: 3541
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-873-367C-163

Alignment Scores:
Pred. No.: 6,36e-186 Length: 3541
Score: 1904.00 Matches: 443
Percent Similarity: 53.66% Conservative: 188
Best Local Similarity: 37.67% Mismatches: 381
Query Match: 32.49% Indels: 164
DB: 13 Gaps: 24

US-09-397-967A-16 (1-1099) x US-09-873-367C-163 (1-3541)
Qy 17 SerLeuSerSerSerGluAlaGlyAlaLeuHisValLeu---LeuProProArgGlyPro 35
Db 121 AACCTGGAGGCGCTCTGAGCCAGGGGTGGAAGTATCTTCTATCTGCGGACAGGAGGCC 180
Qy 36 GlyProProGlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuVal 55
Db 181 -----CTCCGGCTGGCGAGTGTGAGAGTACACAGCAGAGGAAGTGTGATC 225
Qy 56 ArgAlaAlaLysAlaCysGlyLeuProValTyrHisSerLeuPheAlaLeuAlaThr 75
Db 226 AGGGCTGACAGGCAATGCGTATCTCTCTCTTTGTCAACACTCTTGGCCCTGTATGAC 285
Qy 76 GluAspPheSerCysTyrPheProProSerHisIlePheCysIleGluAspValAspThr 95
Db 286 GAGAACACCAAGCTCTGGTATGCTCCAAATCCAGCATCACCGTTCATGACAGATGTC 345
Qy 96 GlnValLeuValTyrArgLeuArgPheTyrPheProAspTyrPheGlyLeu----- 112
Db 346 CTCCTGGCTCCACTACCGGATGAGGTTCTATTTCACCAATGGCATGGAAACCAACGCAAT 405
Qy 113 ---GluThrCysHisArgPheGlyLeuArgLysAspLeu----- 124
Db 406 GAGCAGTCACTGGCGTCAATTCCTCAAGAGACAGAAATGGCTACGAGAAAAAAG 465
Qy 125 -----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis 141
Db 466 ATTCAGATGCAACCCCTCTCTGATGCGAGTCACTGAGATATCTGTTGCTCAGGGA 525
Qy 142 ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----- 159
Db 526 CAGTATGATTGCTGAAATGCTGCTCTCTATTTCGAGACCCCAAGACCCGAGGATGGA 585
Qy 160 -----GlyGluPheLeuSerLeuAlaValLeuAspLeuAlaGlnMetAlaArg 175
Db 586 CATGATATTGAGAACAGTGTCTAGGATGGCTGCTGCTGGCCATCTCACACTATGCCATG 645
Qy 176 GluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrIysAlaCysLeuPro 195
Db 646 ATGAAGAAGATGCGATTGCGCAGAACTGCCCAAGGACATCAGCTTACAAAGCGATATATCCA 705
Qy 196 ProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgGlyLeu----- 213
Db 706 GAAACATTGAATGAATGCTCATCAGACAGAGGAAGCACTTCTCACCAGGATGGGATTAATAT 765
Qy 214 -----ArgArgThrValValLeuAlaLeuLeuPro 223
Db 766 GTTTTCAAGGATTTCCTAAAGGAATTTAAACACAGACCATTTGTGACAGCAGCGTG--- 822
Qy 224 CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlaLysTyrIleLeuAspLeuGlu 243

```

Db 823 -----TCCACGCGATGACCTGGAAGGTGAATACTCTGGCTACCTTGGAA 864  
Qy ArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu 263  
Db 865 ACTTGTGACAAACATTACGGTGCTGAATATTT----- 897  
Qy 264 GluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTyr---SerSerAsnAsp 282  
Db 898 GAGACTTCCATGTTACTGATTTTCACAGAAATGAGATGAATGGTTTCATTCGATGAC 957  
Qy 282 ----- 282  
Db 958 GGTGGAAACGTTCTCTACTACGAAGTGATGGTGAAGTCTTGGAAATCCAGTGGAGG 1017  
Qy 282 ----- 282  
Db 1018 CATAAACCAATGTTGTTCTGTTGMAAAGGAAATAAATAACGAGCGGAAAAAATCG 1077  
Qy 283 -----GluLeuPheGlnThrPhe 288  
Db 1078 GAAATAAAGACAAAGAGATGAGGAGAAAACAAGATCCGGGAAGAGTGGAAACATTTT 1137  
Qy 289 CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAla 308  
Db 1138 TCATCTTCCCTGAAATCACTCAATTTGTAATAAAGGAGTCT----- 1179  
Qy 309 GlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhe 328  
Db 1180 -----GTGGTCAGCATTAACAGCAGGACCAACAGAAATAATGGAATCGAAGCTC 1227  
Qy 329 ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle 348  
Db 1228 TCTTCCACAGGAGGCTTGTCTTGTGTCCTGTGTAGATGCTACTTCCGGCTCACA 1287  
Qy 349 CysAspSerArgHisTyrPheCysLeuValAlaProProArgLeuGluGluGlu 368  
Db 1288 GCAGATGCCCATCAATCACTCTGCACCGAGTGGCCCCCGCTGTGATCGTCCACACATA 1347  
Qy 369 AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisIysLeuLysAlaAla 388  
Db 1348 CAGATGGCTGTATGTGTCATGTCCTGACAGATACCGCATCAATAAATGCGCAAGAA 1407  
Qy 389 GlySerLeuProGlyThrThrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu 408  
Db 1408 GGAAGCGAGGAGGGATGTACGTCTGAGTGGAGCTGCACCGACTTTGACACATCCTC 1467  
Qy 409 LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys 422  
Db 1468 ATGACCGTCACTGCTTTGAGAAGTCTGACAGGTGCGAGGTGCGCCAGAGCAGTTCAAG 1527  
Qy 423 GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro 442  
Db 1528 AACTTTCAGATC---GAGGTGGAGAGGGCGGCTACAGTCTGCAGGTTCCGACCGCAGC 1584  
Qy 443 HisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGly 462  
Db 1585 TTCCCGAGCTTGGAGACCTCATGAGCCACTCAAGAGCAGATCTTGGCCGCGGATAAC 1644  
Qy 463 AlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIle 482  
Db 1645 ATCAGCTTCATGTAAACGCTCTGCGACGCCCAAGCCCGAGAAATCTCCAACCTGCTG 1704  
Qy 483 ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla 502  
Db 1705 GTGGCTACTAAG-----AAAGCCAGGAGTGGCAGGCC---GTCTACCCC 1746  
Qy 503 LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluThrPheGlnAsnLeu 522  
Db 1747 ATGAGCCAGCTGAGTTTCGATCGATCCCTCAAGAAGGATCTGTGCGAGCGGAGCACCTT 1806  
Qy 523 GlyHisGlySerPheThrLysIlePheArggIysSerArgGluValValAspGlyGlu 542

Db 1807 GGGAGAGGCACGAGAACACACATCTATTCTGGGACCTGATGATTAACAGGATGACGAA 1866  
Qy 543 ThrHisAspSerGlu-----ValLeuLysValMetAspSerArgHisArg 558  
Db 1867 GGAACCTTCTGAAGAGAGAGATAAAGTGATCTCTCAAAAGTCTTAGACCCAGCCACAGG 1926  
Qy 559 AsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis 578  
Db 1927 GATATTTCCCTGGGCTCTCTCGAGCGACCCAGCAGCATGTAGACACAGTCTCCACAAAC 1986  
Qy 579 LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe 597  
Db 1987 ATCTGTACCTCTATGGGCTGTGTGCGCAGCTGGAGAAATATCATGGTGAAGAGTTT 2046  
Qy 598 ValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer 617  
Db 2047 GTGGAAGGGGCTCTCTGGATCTCTTCATGCACCGGAAAGCTGATGTCCTTACCACACCA 2106  
Qy 618 TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly 637  
Db 2107 TGGAAATTCAAAGTTGCCAAACAGCTGGCCAGTCCCTGAGCTACTTTGGAGATAAAGAC 2166  
Qy 638 LeuProHisGlyAsnValSerAlaArgLysValLeuAlaArgGluGlyGlyAspGly 657  
Db 2167 CTGCTCCATGGAATGTGTACTTAAACCTCTCTGGCCCGTGGAGGAATCGACAGT 2226  
Qy 658 Asn---ProProPheIleLysLeuSerAspProGlyValSerProThrValLeuSerLeu 676  
Db 2227 GAGTGTGGCCCATTCATCAAGCTCAGTGACCCCGGCATCCCATTCATCGTGTCTCTAGG 2286  
Qy 677 GluMetLeuThrAspArgIleProTyrValAlaProGluCysLeuGlnGluAlaGlnThr 696  
Db 2287 CAGAATGCAATGAACGAATCCCATGATGATGCTCTGAGTGTGTGAGGATCCCAAGAAC 2346  
Qy 697 LeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrThrTrpGluValPheGlnArg 716  
Db 2347 CTGAGTGTGCTGCTGACAAAGTGAGCTTTGGAAACACGCTCTG3GAAATCTGCTACAAT 2406  
Qy 717 GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGln 736  
Db 2407 GGGCAGATCCCTTGAAGACAAGAGCTGATTGAGAAAGAGAGATTCATGAAAGCGG 2466  
Qy 737 GlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAla 756  
Db 2467 TGCAGCCGAGTGACACCATCATGATAAGAGCTGGCTGACCTCATGACCCGCTGCATGAAC 2526  
Qy 757 TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGlyLeuIle 776  
Db 2527 TATGACCCCAATCAGAGGCTTTCTTCGAGGCCATCATGAGACACATTAATAAGCTTGAA 2586  
Qy 777 ThrSerAspTyrGluLeuLeuSer-----AspProThrProGlyIleProSer 792  
Db 2587 GAGCAGAATCCAGATATGTTTCCAGAAAAAACCAGCCAACTGAA----- 2634  
Qy 793 ProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIle 812  
Db 2635 -----GTGGACCCCCACACAT 2649  
Qy 813 PheGluGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerVal 832  
Db 2650 TTTGAGACGGCTTCTTAAGAGGATCCCTGACTTTGGGAGAGGGCCACTTTGGGAAGGTT 2709  
Qy 833 GluLeuCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGln 852  
Db 2710 GAGCTCTGCGAGTATGACCCC---GAAGACATACAGGGGAGCAGGTGGCTGTAAATCT 2766  
Qy 853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnLeu 871  
Db 2767 CTGAAGCCCTGAGATGGAGGTAAACCATAGTGTATCTGAAAAAGAAATCGAGATCTTA 2826  
Qy 872 LysAlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArg 891  
Db 2827 AGGAACCTCTATCATGAGAACATTGTGAAGTACAAAGGAATCTGCACAGAACGCGAGGA 2886

QY	892	GlnSerLeuArgGluValMetGluTyrLeuProSerGlyCysLeuArgAspLeuLeuGln	911
DB	2887	AATGGTATTAAAGCTCATCGTAATTTTCGCCTTCGGGAAGCCTTAAGGAATATCTTCCA	2946
QY	912	ArgHisArgGly--LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysIys	930
DB	2947	AAGAAATGAACAAACAAATAAACCTCAACACGACGCTAAATAATGCGCTTCAGATTTGTAAAG	3006
QY	931	GlyMetGluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaalaArgAsnIle	950
DB	3007	CGGATGACATATTTGGTCTTCGCAATACGTTTACCGGGACATTGGCAGCAGAAATGTC	3066
QY	951	LeuValGluSerGluAlaHisValIysIleAlaAspPheGlyLeuAlaIysLeuLeuPro	970
DB	3067	CTTGTTGAGAGTGAACACCAAGTGAAATTTGGAGACTTCGGTTTAAACCAAGCAATTGAA	3126
QY	971	LeuGlyIysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTyrTyrAla	990
DB	3127	ACCGATAAAGGAGTATTACACCGTCAAGGATGACCGGGACACGCCCTGTGTTTGGTATGCT	3186
QY	991	ProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyVal	1010
DB	3187	CCAGAATGTTTTAATGCAATCAAATTTATATGCCCTCTCAGCTGTGGTCTTTTGGAGTC	3246
QY	1011	ValLeuTyrGluLeuPheThrTyrCysAspIysSerCysSerProSerAlaGluPheLeu	1030
DB	3247	ACTCTGCATGAGCTGCTGACTTACTGTGATTCAGATTCTAGTCCCATGGCTTTGTTCTCTG	3306
QY	1031	ArgMetMetGlyProGluuArgGluGlyProProLeuCysArgLeuLeuGluLeuAla	1050
DB	3307	AAAATGATAGGCCCAACCAATCCATGGCCAGATGACAGCTCAAGACATTGTAATACGTTAAAA	3366
QY	1051	GluGlyArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGln	1070
DB	3367	GAAGGAAACCGCTGCCGTGCCACCTAACTGTCCAGATGAGGTTTATCAGCTTATGAGA	3426
QY	1071	LeuCysTyrPalaProGluProHisAspArgProAlaPheAlaThrLeu	1086
DB	3427	AAATGCTGGGATTTCCCAACCTCCAACTCCGACCAAGCTTTCAGAACCTT	3474

## RESULT 11

```

US-09-962-854A-2
; Sequence 2, Application US/09962854A
; Publication No. US20030044803A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Annette, Soerensen B.
; APPLICANT: Hernandez, Javier Martin
; TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with a
; TITLE OF INVENTION: expression of JAK1
; FILE REFERENCE: A-70020/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,854A
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 03/668,644
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 3538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-854A-2

```

Alignment Scores:		
Pred. No.:	1,628-179	Length: 3538
Score:	1842.00	Matches: 439
Percent Similarity:	53.7%	Conservative: 191
Best Local Similarity:	37.4%	Mismatches: 386
Query Match:	31.43%	Indels: 159
DB:	11	Gaps: 22

US 96-287-867A-16 (1-1099) x US-09-962-854A-2 (1-3538)

Qy	17	SerLeuSerSerSergLunAlaglyAlalauHisValLeu---LeuProProArgGlyPro	35
Db	121	AACTCGAGGCCCCCTGAGCCAGCGGGTGAAGTAGTTCTTCTATCTGTCCGACAGGAGGCC	180
Qy	36	GlyProGrodlnArgLeuSerPheSerPheGlyAspTyrLeuAlaGluAspLeuCysVal	55
Db	181	-----CTCCGGCTGGCAGTGAGAGATACACAGCAGAGAGAACTGTGCATC	225
Qy	56	ArgAlaAlalysalaCysGlyIleLeuProValTyrHisSerLeuPheAlaIleAlaThr	75
Db	226	AGGGCTGCACAGCATGCGGTATCTCTCTCTTTCTCAACCTCTTTGGCCTGTATGAC	285
Qy	76	GluAspPheSerCysTrpPheProPserHisIlePheCysIleGluAspValAspThr	95
Db	286	GAGAACAACCAGCTCTGGTATGCTCCAATTCCACCATCACCGTTGATGACAAAGATGCC	345
Qy	96	GlnValLeuValTyrArgLeuArgPheTyrPheProAspTrpPheGlyLeu---	112
Db	346	CTCCGGCTCCACTACCGATGAGTTCTATTTCACCAATTGGCATGGAACCAACGACAAT	405
Qy	113	---GluThrCysHisArgPheGlyLeuArgLyAspLeu-----	124
Db	406	GAGCAGTCAGTGGCGGTCAATTCTCCAAAAGAACGAGAAAAATGGCTACGAGAAAAAAG	465
Qy	125	-----ThrSerAlaIleLeuAspLeuHisValLeuGluHisLeuPheAlaGlnHis	141
Db	466	ATTCACATGCAACCCCTCTCTTGATGCAGCTCACTGGAGTATCTGTTGCTCAGGGA	525
Qy	142	ArgSerAspLeuValSerGlyArgLeuProValGlyLeuSerMetLysGluGln----	159
Db	526	CAGTATGATTGGTGAATCCCTGGCTCTATTTCGAGACCCCAGAACCCGACGAGATGA	585
Qy	160	-----GlyGluPheLeuSerLeuAlavalLeuAspLeuAlaGlnMetalArg	175
Db	586	CATGATATTGAGAACGAGTGTCTAGGATGGCTGTCTGTGCCATCTCACACTATGCCATG	645
Qy	176	GluGlnAlaGlnArgProGlyGluLeuLeuLysThrValSerTyrLysAlaCysLeuPro	195
Db	646	ATGAAGAAGATGCAGTTGCCAGAACTGCCCAAGCATCAG-GTAAGGCGATATATCCA	704
Qy	196	ProSerLeuArgAspValIleGlnGlyGlnAsnPheValThrArgArgile-----	213
Db	705	GAACATTGTAATAAGTCCATCAGACAGAAACCTTCTACCCAGGATGGGTAATAATAT	764
Qy	214	-----ArgArgThrValValLeuAlaLeuLeuPro	223
Db	765	GTITTCAGGATTTCTTAAGGAATTTAAACAAGACCATTTGTGCAGACGCGTG---	821
Qy	224	CysGlyArgLeuProGlyArgProTyrAlaLeuMetAlalysTyrIleLeuAspLeuGlu	243
Db	822	-----TCCACGCGATGCTGAAGGTGAATTAATCTTGGCTACCTTGGAA	863
Qy	244	ArgLeuHisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu	263
Db	864	ACTTTTGACAAAACATTACGGTGTGTAATATT-----	896
Qy	264	GluProGlyLeuLeuArgValAlaGlyAspAsnGlyIleProTrp---SerSerAsnAsp	282
Db	897	GAGACTTCCATGTTACTGATTTTCATCAGAAAATGAGATGAATTGGTTTCATTCGAATGAC	956
Qy	282	-----	282
Db	957	GGTGGAAAGCTTCTACTACAAAGTGAATGGTGAATCTGGGAATCTTGAATCCAGTGAGG	1016
Qy	282	-----	282
Db	1017	CATAAACCAATGTTGTTCTGTTCAAAAGGAAAAATAAATAACTGAAGCGGAAAAAACTG	1076
Qy	283	-----GluLeuPheGlnThrPhe	288
Db	1077	GAATAATAACAAGAGATGAGAGAAAAACAAGATCCGGGAAGTAGTGAAACAATTTT	1136

Qy	289	CysAspPheProGluIleValAspValSerIleAsnGlnAlaProArgValGlyProAla	308
Db	1137	TCTTACTTCCTCGAATCACTCACTATGTTAATAAGGAGTCT	1178
Qy	309	GlyGluHisArgLeuValThrValThrArgMetAspGlyHisIleLeuGluAlaGluPhe	328
Db	1179	-----GTGGTTCAGCATTAACAGCAGACCAACAGAAAATGGAACTGAAGCTC	1226
Qy	329	ProGlyLeuProGluAlaLeuSerPheValAlaLeuValAspGlyTyrPheArgLeuIle	348
Db	1227	TCTTCCACAGGAGGCGCTTGCTTGTGTGCTGTAGATGCTACTTCCGGCTCACA	1286
Qy	349	CysAspSerArgHisTyrPheCysLysGluValAlaProProArgLeuLeuGluGlu	368
Db	1287	GCAGATCCCATCAATACCTCTGCACGCGGTGGCCCCCGTTGTATGCTCCACAACATA	1346
Qy	369	AlaAspValCysHisGlyProIleThrLeuAspPheAlaIleHisLysLeuIleAlaAla	388
Db	1347	CAGAAATGGCTGTCTATGTGTCAATCTGTACAGATACGCCATCAATAAATTCGGCAAGAA	1406
Qy	389	GlySerLeuProGlyThrTyrIleLeuArgArgSerProGlnAspTyrAspSerPheLeu	408
Db	1407	GGAAAGCAGGAGGCGATGTACGTCTCAGGTGGGCTGC-ACCGACITTTGACAACTCTC	1465
Qy	409	LeuThr---AlaCysValGlnThrPro-----LeuGlyProAspTyrLys	422
Db	1466	ATGACCTGCTACCTGCTTTGAGAAAGTCTGACAGGTGCGAGGTGCCAAGACGATTTCAAG	1525
Qy	423	GlyCysLeuIleArgGlnAspProSerGlyAlaPheSerLeuValGlyLeuSerGlnPro	442
Db	1526	AACTTTCAGATC--GAGGTGCAGAGGGCGGTACAGTCTGCAGGTTTCGGACCGCAGC	1582
Qy	443	HisArgSerLeuArgGluLeuLeuAlaAlaCysTrpAsnSerGlyLeuArgValAspGly	462
Db	1583	TTCCCCAGCTTGGAGACCTCATGAGCCACTCAAGAGCAGATCTCTGCCCGGATAAC	1642
Qy	463	AlaAlaLeuTyrLeuThrSerCysCysAlaProArgProLysGluLysSerAsnLeuIle	482
Db	1643	ATCAGCTTCATGTATAAACGCTGTGCAGCCCAAGCCCCGAGAAATCTCCAACCTGCTG	1702
Qy	483	ValValArgArgGlyCysAsnProAlaProAlaProGlyCysSerProSerCysCysAla	502
Db	1703	GTGGCTACTAAG-----AAAGCCCCAGGAGTGGCAGCCCC--GTCTACCCC	1744
Qy	503	LeuThrGlnLeuSerPheHisThrIleProThrAspSerLeuGluTrpHisGluAsnLeu	522
Db	1745	ATGAGCCAGCTGAGTTTCGATCGGATCTCAAGNAGGATCTGGTGCAGGCGGAGACCTT	1804
Qy	523	GlyHisGlySerPheThrLysIlePheArgGlySerArgArgGluValValAspGlyGlu	542
Db	1805	GGGAGAGGCACGAGAACACACATCTATTCTGGGACCCCTGATGGATACAAGGATACGAA	1864
Qy	543	ThrHisAspSerGlu-----ValLeuLeuLysValMetAspSerArgHisArg	558
Db	1865	GGAACTTCTCAAGAGAAAGATAAAGTGATCTCTAAAGTCTTAGACCCCGCCACAGG	1924
Qy	559	AsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnValSerTyrProHis	578
Db	1925	GATATTTCCTGCCCTTCTTCGAGGCGACGAGATGATGAGACAGGTCTCCCCAACAC	1984
Qy	579	LeuValLeuLeuHisGlyValCysMetAlaGly---AspSerIleMetValGlnGluPhe	597
Db	1985	ATCGTGTACCTCTATGGCGTCTGTGTCCGCGACGTGGAGAATATCATGTGTGAAGAGTTT	2044
Qy	598	ValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuValSerAlaSer	617
Db	2045	GTGAAGGGGGTCTCTGGATCTCTTCATGCACCGGAAAAGCGATGCTCTTACCACCA	2104
Qy	618	TrpLysLeuGlnValThrLysGlnLeuAlaTyrAlaLeuAsnTyrLeuGluAspLysGly	637
Db	2105	TGGAAATTCAAAGTTGCCAAACAGCTGGCCAGTGGCCTGAGCTTACTTGGAGGATAAAGAC	2164
Qy	638	LeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGlyGlyAspGly	657

Db	2165	CTGGTCCATGAAATGTGTGTACTATAAAACCTCTCTCTGGGCCCTGAGGCGCATCGACAGT	2224
Qy	658	Asn---ProProPheIleLysSerAspProGlyValSerProThrValLeuSerLeu	676
Db	2225	GAGTGTGGCCCGTTCATCAAGCTCAGTGACCCCGGCATCCCATTAACGGTGTCTGTCTAGG	2284
Qy	677	GluMetLeuThrAspArgIleProTrrpValAlaProGluCysLeuGlnGluAlaGlnThr	696
Db	2285	CAAGAATGCATTTGAACGAATCCCATGGATGTCTCTGAGTGTGTGAGGACTCCAGAAGC	2344
Qy	697	LeuCysLeuGluAlaAspLysTrrpGlyPheGlyAlaThrThrTrrpGluValPheGlnArg	716
Db	2345	CTGAGTGTGGTGTGTGCAAGTGAAGCTTTGGAAACCAACCGCTCTGGGAAATCTGTACAAT	2404
Qy	717	GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGluAspGln	736
Db	2405	GGCAGATCCCTTTGAAAGACAAGACGCTGATTGAAAGAGAGATCTTATGAAAGCCGG	2464
Qy	737	GlyGlnLeuProAlaLeuLysTrrpThrGluLeuAlaGlyLeuIleThrGlnCysMetAla	756
Db	2465	TGCAGGCCAGTGACACCATCATGATGAAGAGCTGGCTGACCTCATGACCCGCTGCATGAAC	2524
Qy	757	TyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAenGlyLeuIle	776
Db	2525	TATGACCCCAATCAGAGCGCTTCTCCGAGCCATCATGAGACATTAATAGCTTGAA	2584
Qy	777	ThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArgAspGlu	796
Db	2585	GAGCAGAAATCCAGATATTGTTTCAGAAAAAACCAACTGAA	2629
Qy	797	LeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArg	816
Db	2630	-----GTGGACCCCAACACATTTTGAAGACGC	2656
Qy	817	HisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeuCysArg	836
Db	2657	TTCCTAAGAGGATCCGTGACTTTGGGAGAGGGCCACTTTGGGAAGGTGAGCTCTGCAGG	2716
Qy	837	TyrAspProLeuGlyAspAenThrGlyProLeuValAlaValLysGlnLeuGln---	855
Db	2717	TATGACCCCAAGGGGCAATACAGGGGACAGGTGGCTGTTAAATCTCTGAAGCCTGAG	2776
Qy	856	SerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeuLysAlaLeuHis	875
Db	2777	AGTGGAGGTAACACATAGCTGATCTGAAAGAGMAATCGAGATCTTAAGAACCTCTAT	2836
Qy	876	SerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerLeuArg	895
Db	2837	CATGAGAACATTTGTAAGTAGTACAAGGAATCTG-CACAGAAGACAGGAATGTTATTAAG	2895
Qy	896	LeuValMetGluTyrLeuProSerGlyCysLeuArgAspLeuGlnArgHisArgGly	915
Db	2896	CTCATCATGAAATTTCTGCCCTTCGGAGACCTTAAGGAATATCTTCCAAGTAATAAGAC	2955
Qy	916	---LeuHisThrAspArgLeuLeuLeuPheAlaTrrpGlnIleCysLysGlyMetGluTyr	934
Db	2956	AAAATAAACCTCAACAGCAGCTAAAAATATGCCGTTTCAGATTGTGAAGGGATGGACTAT	3015
Qy	935	LeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSer	954
Db	3016	TTGGGTTCGCGCAATACGTTCAACCGGACTTGGCAGCAAGAAATGCTCCTGTTCAGAGT	3075
Qy	955	GluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGlyLysAsp	974
Db	3076	GAACCAAGTGAATTTGGAGACTTCGGTTTTAACCAAGCAATTTGAACCCGATAGGAG	3135
Qy	975	TyrTyrValValArgGluProGlyGlnSerProIlePheTrrpTyrAlaProGluSerLeu	994
Db	3136	TATTACACCGTCAAGGATGACCGGGACAGCCCTGTGTTTGGTATGCTCCAGAATGTTTA	3195
Qy	995	SerAspAsnIlePheSerArgGlnSerAspValTrrpSerPheGlyValValLeuTyrGlu	1014







Db	3760	TGGGAGACAGAGGCGTCTTTTCGCCCAACCTTCGAGAACCTCATACCCATTCTG	3813
RESULT 13			
US-10-062-674-2215			
; Sequence 2215, Application US/10062674			
; Publication No. US20040005559A1			
; GENERAL INFORMATION:			
; APPLICANT: Loring, Jeanne F.; Kaser, Matthew R.			
; TITLE OF INVENTION: MARKERS OF NEURONAL DIFFERENTIATION AND MORPHOGENESIS			
; FILE REFERENCE: PA-0026-1 CIP			
; CURRENT APPLICATION NUMBER: US/10/062,674			
; PRIORITY FILING DATE: 2002-01-30			
; PRIOR APPLICATION NUMBER: US 09/625,102			
; PRIORITY FILING DATE: 2000-07-24			
; NUMBER OF SEQ ID NOS: 2217			
; SOFTWARE: PERL Program			
; SEQ ID NO 2215			
; LENGTH: 5228			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc feature			
; OTHER INFORMATION: Incyte ID No. US20040005559A1 903876.6			
; FEATURE:			
; NAME/KEY: unsure			
; LOCATION: (1) ... (5228)			
; OTHER INFORMATION: a, t, c, g, or other			
US-10-062-674-2215			
Alignment Scores:			
Pred. No.:	2,15e-138	Length:	5228
Score:	1416.50	Matches:	451
Percent Similarity:	51.35%	Conservative:	176
Best Local Similarity:	36.94%	Mismatches:	393
Query Match:	24.68%	Indels:	203
DB:	12	Gaps:	31
US-09-397-967A-16 (1-1099) x US-10-062-674-2215 (1-5228)			
QY	7	GluThrProLeuIleProGlnArgSerCysSerLeuSerSerSerGluAlaGlyAlaLeu	26
Db	370	GAGGCCCTTGATGCGCCAGGGGNTGGCAAGCTGGATCCTTCG	414
QY	27	HisValLeuLeuProProArgGlyProGlyProGlnArgLeuSerPheSerPheGly	46
Db	415	---TNAATCTGTCGGCGCACAGGGAAGCCCTCCG	462
QY	47	AspTyrLeuAlaGluAspLeuCysValArgAlaAlaLysAla	65
Db	463	GAGTACACACGACAGGAACCTGTCATCAGGGCGGTGCACAGGGCATGCCGGTATCTCTCC	522
QY	65	oValTyHisSerLeu---PheAlaLeuAlaThrGluAspPheSerCysTrpPheProPr	84
Db	523	TTCTTTGTCAACTCTCTTTTGGCTGTATGACGAGACACCAAGCTCTGGTATCTCC	582
QY	84	o--SerHisIlePheCysIleGluAspValAspThrGlnValLeuValTyrArgLeuAr	103
Db	583	CAAACTCGCAATCACCGTTGATGACAAAGATGTCTCCGCTCCACTACCGTCCACCG	642
QY	103	gPheTyrPheProAspTrpPheGlyLeuGluThr	115
Db	643	GTTCATATTCACCAATTCGGATGGAA--CACAACCAATGACATCATGTGGCGTCAT	701
QY	116	-----HisArgPheGlyLeuArgLysAspLeuThrSerAlaLle	128
Db	702	TCTCCCAAGAGCCAGAAATTTGGTACCGAGAAAAAAGATTCCAGATGCAACCCCT	761
QY	129	----LeuAspLeuHisValLeuGluHisLeuPheAlaGlnHisArgSerAspLeuValse	147
Db	762	CTCGCTTGATGCCAGCTCACTGGAGTATCTGTCTCAGGACAGTAGATTTGGTGA	821
QY	147	rglyArgLeuProValGlyLeuSerMetLysGluGln	161

Db	822	ATGCTGCTCTCTATTTGAGACCCCAAGACGAGGAGATGGACATGATATTGAGACGA	881
QY	161	upheLeuSerLeuAlaVal-LeuAspLeuAlaGlnMetAlaArgGluGlnAlaGlnArgp	181
Db	882	GTGTCTAGGATGGCTGTCGTGGCCATCTCACATATGCCATGATGAAGAGATGCAGT	941
QY	181	roGlyGluLeuLeuLys-ThrValSerTyrIysAlaCysLeuProProSerLeuArgasp	200
Db	942	TGCCAGTACTGCCAAGGACATCAAGCTACAAGCGATATATTCAGAAACATTTGAATAAG	1001
QY	201	ValIleGlnGlyGlnAspPheValThrArgArgArgille	213
Db	1002	TCCATCAGACAGAGGAACCTTCTCCAGGATGCGGATAAATAATGTTTTCAAGATTTTC	1061
QY	214	-----ArgArgThrValValLeuAlaLeuLeuProCysGlyArgLeu	227
Db	1062	CTAAAGGAATTTAAACAACAGACCTTT-GTGACAGACAGCGCTG	1105
QY	228	ProGlyArgProTyrAlaLeuMetAlaIysTyrIleLeuAspLeuGluArgLeu	245
Db	1106	-----TCCACGATGACCTGAAGGTGAATATCTTGGCTACCTTGGAACTTTGACAAA	1159
QY	246	---HisProAlaAlaThrThrGluThrPheArgValGlyLeuProGlyAlaGlnGlu	264
Db	1160	CATTACGCTGCTGAATATTTGAGACTTCCATGTTACTGATTTTCATCAGAAATGAGATG	1219
QY	265	Pro-GlyLeuLeuAlaGlyAlaLala	272
Db	1220	AATTGTTTTCATTCGATGACGGTGAAACGTTCTCTACTACGAGTGTGTGCTGGG	1279
QY	272	YAspAsnGlyIleProTyrSerSerAsnAsp	282
Db	1280	GAATCTTGAATCCAGTGGAGGCATAAACCAAAATGTTTCTTCTGTTGAAAGAAAAA	1339
QY	282	-----	282
Db	1340	TAAACTGAAGCGGAAAAAATCTGAAATAAACAAGAGGATGAGGAAAAACAAGAT	1399
QY	283	---GluLeuPheGlnThrPheCysAspPheProGluIleValAspValSerIleAsnG	301
Db	1400	COGGAGAGAGTGAAACAATTTTCTTCTTCTCCCTGAATCACTCACATTTGTAATAAGGA	1459
QY	301	nAlaProArgValGlyProAlaGlyGluHisArgLeuValThrValThrArgMetAspG	321
Db	1460	GTCT-----GTGGTCAGCATTAACAAGCAGGACAA	1489
QY	321	yHisIleLeuGluAlaGluPheProGlyLeuProGluAlaLeuSerPheValAlaLeuVa	341
Db	1490	CAAGAAATGGAATGAAGCTCTTCTCCACGAGGAGGCTTGTCTTTGTGTTCTTGGT	1549
QY	341	lAspGlyTyrPheArgLeuIleCysAspSerArg-HisTyrPheCysLysGluValAla	361
Db	1550	AGATGGCTACTTCGGCTCACAGAGATGCCATCTATTACCTCTGCACCGACGTCGCC	1609
QY	361	roProArgLeuLeuGluGluAlaAspValCysHisGlyProIleThrLeuAspPheA	381
Db	1610	CCCCGTGTGATCGTCCCAACAACATACAGATGGCTCATGTGTCTCAATCTGTACAGAATCG	1669
QY	381	lalleHisLysLeuLysAlaAlaGlySerLeuProGlyThrTyrIleLeuArgArgSerp	401
Db	1670	CCATCAATAATTCGGCAAGAGGAGGAGGAGGATGCTGCTGTAGGTGGAGCT	1729
QY	401	roGlnAspTyrAspSerPheLeu-LeuThrAla---CysValGlnThrPro	416
Db	1730	GCACCGACTTGTGACAACTCTCAATGACCGCTCAACTGCTTGTGAGAGCTCTGAGCAGGTG	1789
QY	417	-----LeuGlyProAspTyrLysGlyCysLeuIleArgGlnAspProSerGlyAlaPhe	434
Db	1790	CAGGTGCGCCAGAGAGCATTTCAAGAACTTTCAGATC---GAGGTGAGAGGCGCGCTAC	1846
QY	435	SerLeuValGlyLeu-SerGlnProHisArgSerLeuArgGluLeuAlaAlaCysTr	454



Db 1847 AGTCTGCAGGCTCTCGGACCGGAGCTTCCCGACATGGGAGACCTCATGAGCCACCTCAA 1906  
Qy 454 pAsnSerGlyLeuArgValAspGlyAlaAlaLeuTyrLeuThrSerCysCysAlaProAr 474  
Db 1907 GAAGCAGATCTTCGCGACGGATAACATCAGCTTCATGCTTAAACCGTGTGCGAGCCCAA 1966  
Qy 474 gProLysGluLysSerAsnLeuLeuValValArgArgGlyCysAsnProAlaProAlaPr 494  
Db 1967 GCCCGAGAAATCTCCAACTCTGTGTGCTACTAAG-----AAAGCCCA 2011  
Qy 494 oGlyCysSerProSerCysCysAlaLeuThrGlnLeuSerPheHisThrIleProThrAs 514  
Db 2012 GGATGGGAGCC-----GTCTACCCCATGAGCGAGCTGAGTTTCGATCGCATCTCAAGAA 2068  
Qy 514 pSerLeuGluTyrHisGluAsnLeuGlyHisGlySerPheThrLysIlePheArgGlySe 534  
Db 2069 GGATCTGTGCGAGCGAGCACCTTGGGAGAGCGACGAGAACACATCTATTCTTGGGAC 2128  
Qy 534 rArgArgGluValValAspGlyGluThrHisAspSerGlu-----ValLeuLe 550  
Db 2129 CCTGATGATTACAGGATGACGAAGAACTTCTGAAGAGAAAGATAAAAGTGATCTT 2188  
Qy 550 uLysValMetAspSerArgHisArgAsnCysMetGluSerPheLeuGluAla-AlaSerL 570  
Db 2189 CAAGTCTTAGACCCAGCCAGCAGGATATTTCCTGGCTTCTTCGAGGGAAGCCAGCA 2248  
Qy 570 eMetSerGlnValSerTyrProHisLeuVal-LeuLeuHisGlyVal-CysMetAlaGl 589  
Db 2249 TGATGAGACAGTCTCCACAAACACATCGTGNCCCTCTATGGCTCTTGTGTCGCGCA 2308  
Qy 589 y---AspSerIleMetValGlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuAr 608  
Db 2309 CGTGAGAAATATCATGTGTGAAGAGTTGTGAAGGGGGTCTCTCTGGATCTCTTCATGCA 2368  
Qy 608 gLysArgGlyHisLeuValSerAlaSerTyrLysLeuGln-ValThrLysGlnLeuAlaT 628  
Db 2369 CCGGAAGCGATGTCTTTACCAACCATTGCGAAATTCAAAGTCGCGAAACAGCTGGCCA 2428  
Qy 628 yAlaLeuAsnTyrLeuGluAspLysGlyLeuProHis-GlyAsnValSerAlaArgLys 647  
Db 2429 GTGCCCTGAGCTACTTGAGGATTAAGACCTGGTCCATTTGGAATGTGTACTATAAAAC 2488  
Qy 648 Val-LeuLeuAlaArgGluGly-GlyAspGlyAsn---ProPro-PheIleLysLeuSer 665  
Db 2489 CTTCCTCTGCGCGTGGGCGATCCGACAGTGTGGGCCCATTCATCAAGCTCAGT 2548  
Qy 666 AspProGlyValSerProThrValLeuSerLeuGluMetLeuThrAspArgIleProTyr 685  
Db 2549 GACCCCGCATCCCATTCAGCGTCTGTCTAGGCAAGATGCAATTGAACGAATCCCATGG 2608  
Qy 686 ValAlaProGluCys-----LeuGlnGluAlaGlnThrLeuCysLeuGluAlaAsp 702  
Db 2609 ATTGCTCTGAGTGTGTGAGGACTCCAGAACCTGAGTGGCTGCGTG-----ACA 2662  
Qy 703 LysTyrPheGlyAlaThrTyrTyrPheGlnArgGlyProAlaHisIleThr 722  
Db 2663 AGTGGAGCTTTGGAAACCGCTCTGGGAAATCTGTACAAATGCGAGATCCCTTGAAA 2722  
Qy 723 SerLeuGluPro-AlaLysLysLeuLysPheTyrGluAspGlnGlyGlnLeuProAlaLe 742  
Db 2723 AGCAGAGCCGCTGATTGAGAAAGAGAGATTCTATGAAGCGCGTGGCAGCGCAGTGACACC 2782  
Qy 742 uLysTyrThrGluLeuAlaGlyLeuIleThrGlnCysMetAlaTyrAspProGlyArgAr 762  
Db 2783 ATCATGTAAAGAGCTGCTGACCTCATGACCCGCTGATCACTATGACCCCAATCAGAG 2842  
Qy 762 gProSerPheAlaIleLeuArgAspLeuAsnGlyLeuIleThrSerAspTyrGluLe 782  
Db 2843 GCCTTTCTCCGAGCCATCATGAGACATTAATAAGCTTGAAGACAGAAATCCAGATAT 2902  
Qy 782 uLeuSer-----AspProThrProGlyIleProSerProArgAspGluLeuCy 798  
Db 2903 TGTTCCAGAAAAAACCAGCCAACTGAA----- 2933

Qy 798 sValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGluGluArgHisLe 818  
Db 2934 -----GTGACCCACACATTTTGAAGCGCTCTCT 2965  
Qy 818 uLysTyrIleSerLeuLeuGlyLysGlyAsn-PheGlySerValGluLeuCysArgTyrA 838  
Db 2966 AAAGAGGATCCGTGACTTGGGAGAGCGCCACCTTTGGGAAGGTTGAGCTCTCGAGTATG 3025  
Qy 838 sProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln----- 854  
Db 3026 ACCCCGAAGGACAAATACAGGGAGCAGGTGCTGTTAAATCTCTGGAGCCTGAGACT 3085  
Qy 855 -----HisSerValProAspGlnGlnArgAspPheGlnArgGluLeuGlnIleL 871  
Db 3086 GGAGGTAACCATAGC-----TGATCTGCAAAAAGGAAATCGAGGATC 3130  
Qy 871 eUlys-----AlaLeuHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyP 889  
Db 3131 TTAAGGGAACTCTATTCTATGAGAAC-----ATTGGAAGTACAAAGGAATCTGTACAGAA 3187  
Qy 889 rGlyArgGln--SerLeuArgLeuValMetGluTyrLeuProSerGlyCysLeuArgAs 908  
Db 3188 GACGAGGAAATGGTATTAAAGCTCATCATGGAATTTCTGCTTCGGAAGCCTTAAGA 3247  
Qy 908 pLeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuLeuPheAlaTyrPl 927  
Db 3248 ATATCTTCCAAAGAAATAGAAACAAATAAAGCTCAACAGCAGCTAAAAATATGCCGTTCA 3307  
Qy 927 nileCysLysGlyMetGluTyrLeuGlyValArgArgCysValHisArgAspLeuAlaAl 947  
Db 3308 GATTTGAAGGGATGAGCTATTGGGTCTCGGCAATACGTTTCCCGGAGCTTGGCAGC 3367  
Qy 947 aArgAsnIle-LeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaL 967  
Db 3368 AAGAAATGTCTCTGTTGAGAGTGAACACCAAGTGAAATTTGGAGACTTCGTTTAAACA 3427  
Qy 967 ysLeuLeuProLeuGlyLysAspTyrTyrValVal-ArgGluProGlyGlnSerProile 986  
Db 3428 AAGCAATTTGAACCCGATGAAGAGTATTACACCGTACAAGGATGACCGGACAGCCCTGTG 3487  
Qy 987 PheTyrTyrAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSer-AspValTr 1006  
Db 3488 TTTTGGTATGCTCCAGATGTTTAAATGCAATCTAAATTTTATAATGCTCTCTGACGCTG 3547  
Qy 1006 p-SerPheGlyValValLeuTyrGluLeu-PheThrTyrCysAspLysSerCysSerPro 1025  
Db 3548 GGTCTTTGGAGTCACTCTGCATGAGCTGGCTGACTTACTGTGATTGAGATTTAGTCCC 3607  
Qy 1026 SerAlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeu 1045  
Db 3608 ATGGCTTTGTTCTGAAAAATGATAGGCCCAACCCATGCGCCAGATGACAGTCAAGACTT 3667  
Qy 1046 LeuGluLeuLeuAlaGluGlyArgArgLeuProProProPro-ThrCysProThrGluVa 1065  
Db 3668 GTGAATACGTTAAAGAGGAAACCGCTCCCGTCCACCTAACTGATCCAGATGAGGT 3727  
Qy 1065 lGlnGluLeuMetGlnLeuCys-TripAlaProGluProHisAspArgProAlaPheAlaT 1085  
Db 3728 TGATCAACTATTGAGGAAATGCTTTGGGAATTCACCACTCAATCGGACGAAGCTTTTCA 3787  
Qy 1085 hrLeu 1086  
Db 3788 ACCTT 3792

RESULT 14  
US-10-205-219-2  
; Sequence 2, Application US/10205219  
; Publication No. US20030138803A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair



```

Db 1729 GAAGTAAAGACCTCGTTCATGGAATGTGTGCACTAAACCTCCTCTGCGCCGTGAG 1788
Qy 654 GlyIyAspGlyAsn---ProPheIleLysLeuSerAspProGlyValSerProThr 672
Db 1789 GCATTCAGTGAATGAGCCCGTTCATAGCTTAGTGAAGCTTGGCATCCAGTCTCT 1848
Qy 673 ValLeuSerLeuGluMetLeuThrAspArgIleProTTPValAlaProGluCysLeuGln 692
Db 1849 GTGCTGACGAGCAAGAGTGCATAGAGCAATCCCTGGATCGCTCGTGTGTGAA 1908
Qy 693 GluAlaGlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGlu 712
Db 1909 GACTCAAGACCTGAGTGTGCTCTGCAAGTGGAGCTTTGGACACCACTCTGGAA 1968
Qy 713 ValPheGlnArgGlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPhe 732
Db 1969 ATCCGCTACGATGGCGAGATCCCATCAAGGACAAGACCTCATTTGAGAAAGAGGTT 2028
Qy 733 TyrGluAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThr 752
Db 2029 TATGAAAGCCGTGCGAGCCAGTGCACCATCTTCAAGGAGCTAGCTGACCTCATGACT 2088
Qy 753 GlnCysMetAlaTyAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeu 772
Db 2089 CGTCGATGAATGATATCCCAACAGAGACCTTTCTCCGAGCCATCATGAGGACAT 2148
Qy 773 AsnGlyLeuIleThrSerAspTyTrpGluLeuSerAspProThrProGlyIleProSer 792
Db 2149 AACAGCTGGAGGAGCAATCCAGACATGTTTCAGAAAGACCCCA----- 2196
Qy 793 ProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyAlaCysGlnAspProAlaIle 812
Db 2197 -----ATAACAGAGGTGATCCCACTCAT 2220
Qy 813 PheGluGlnArgHisLeuLysTyxIleSerLeuLeuGlyLysGlyAsnPheGlySerVal 832
Db 2221 TTTGAAAGCGTTTCCATAAGAGGATTCGTGACTTTGGAGAGGTCACCTTTGGGAAGTT 2280
Qy 833 GluLeuCysArgTyAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGln 852
Db 2281 GAGCTTCGAGATATGATCTTCAGGGAGAACACAGGGGAACAGGTAGCTGTCAAGTCC 2340
Qy 853 LeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGluIleGlnIleLeu 871
Db 2341 CTGAAGCTCGAGCTGAGGTGAGGTAACACATGCTGATCTGAAGAGGAGATAGATCTTA 2400
Qy 872 LysAlaLeuHisSerAspPheIleValLysTyArgGlyValSerTyGlyProGlyArg 891
Db 2401 CGGAACCTCTACCACGAGAACATTTGTAAGTATAAAGGAATCTGCATGGAAGACGGGG 2460
Qy 892 GlnSerLeuArgLeuValMetGluTyTrpLeuProSerGlyCysLeuArgAspLeuGln 911
Db 2461 AATGGTATCAAGCTCATCATGAGTTCGTGCTTCGGGAAGCCCTAAAGGAATATCTGCCA 2520
Qy 912 ArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLys 930
Db 2521 AAGAATAAGAACAAATACTCAACCTCAACAGCAGCTAAATAATGCCATCCAGATTTGTAG 2580
Qy 931 GlyMetGluTyTrpLeu 935
Db 2581 GGGATGGACTATCTCG 2595

```

RESULT 15

US-10-131-410-22

; Sequence 22, Application US/10131410

; Publication No. US20030235915A1

; GENERAL INFORMATION:

; APPLICANT: SPECHT, THOMAS

; APPLICANT: HINZMANN, BERND

; APPLICANT: SCHMITT, ARMIN

; APPLICANT: PILARSKY, CHRISTIAN

; APPLICANT: DAHL, EDGAR

; APPLICANT: ROSENTHAL, ANDRE

; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST

; FILE OF INVENTION: TUMORS

; FILE REFERENCE: SCH-1763

; CURRENT APPLICATION NUMBER: US/10/131,410

; CURRENT FILING DATE: 2002-04-25

; PRIOR APPLICATION NUMBER: 09/646,673

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: PCT/DE99/00908

; PRIOR FILING DATE: 1999-03-19

; NUMBER OF SEQ ID NOS: 202

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 22

; LENGTH: 2327

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-131-410-22

Alignment Scores:

Pred. No.:	1,64e-66	Length:	2327
Score:	746.00	Matches:	151
Percent Similarity:	64.06%	Conservative:	54
Best Local Similarity:	47.19%	Mismatches:	97
Query Match:	12.73%	Indels:	18
DB:	12	Gaps:	3

US-09-397-967A-16 (1-1099) x US-10-131-410-22 (1-2327)

Qy 769 LeuArgAspLeuAsnGlyLeuIleThrSerAspTyTrpGluLeuLeuSerAspProThrPro 788

Db 8 GTCCGAGACATTAATAGCTTTGAAGAGCAGATCCAGATATTGTTTCAGAAAAAACCA 67

Qy 789 GlyIleProSerProArgAspGluLeuCysValAlaGlyAlaGlnLeuTyAlaCysGln 808

Db 68 GCAACTGAA-----GTG 79

Qy 809 AspProAlaIlePheGluGluArgHisLeuLysTyxIleSerLeuLeuGlyLysGlyAsn 828

Db 80 GACCCACACATTTTGAAGAGCGCTTCTTAAAGAGAGTCCGTGACTTTGGGAGAGGCCAC 139

Qy 829 PheGlySerValGluLeuCysArgTyAspProLeuGlyAspAsnThrGlyProLeuVal 848

Db 140 TTTGGGAAGGTTGAGCTCTGAGGTATGACCCCGAAGGGACCAATACAGGGGAGCAGGTG 199

Qy 849 AlaValLysGlnLeuGln---HisSerValProAspGlnGlnArgAspPheGlnArgGlu 867

Db 200 GCTCTTAAATCTCTGAAGCGCTGAGAGTGGAGGTAACCAACATAGCTGATCTGAAAAAGAA 259

Qy 868 IleGlnIleLeuLysAlaLeuHisSerAspPheIleValLysTyArgGlyValSerTy 887

Db 260 ATCGAGATCTTAAGGAACCTCTATCATGGAACATTTGAAGTACAAAGGAATCTGCACA 319

Qy 888 GlyProGlyArgGlnSerLeuArgLeuValMetGluTyTrpLeuProSerGlyCysLeuArg 907

Db 320 GAAGACGGAGGAAATGCTATTAAAGCTCATCATGGAATTTCTGCCCTTCGGAAGCCTTAAG 379

Qy 908 AspLeuLeuGlnArgHisArgGly---LeuHisThrAspArgLeuLeuPheAlaTrp 925

Db 380 GAATATCTTCAAGAAATAAGAAACAAATAAACCTCAACAGCAGCTTAAATAATATGCCGT 439

Qy 927 GlnIleCysLysGlyMetGluTyTrpLeuGlyAlaArgArgCysValHisArgAspLeuAla 946

Db 440 CAGATTGTAAAGGGATGGACTATTGGGTCTTCGGCAATACGCTTACCCTGGGACTTGGCA 499

Qy 947 AlaArgAsnIleLeuValGluSerGluAlaHisValLysIleAlaAspPheGlyLeuAla 966

Db 500 GCAAGAAATGCTCTTGTGAGAGTGAACCAAGTGAATAATTTGAGACTTCGGTTTAACC 559

Qy 967 LysLeuLeuProLeuGlyLysAspTyTrpValValArgGluProGlyGlnSerProIle 986

Db 560 AAAGCAATTGAAACCGATAGGAGTATTACCGCTCAAGGATGATCCCGGACACGCCCTGTG 619

Qy 987 PheTrpTyAlaProGluSerLeuSerAspAsnIlePheSerArgGlnSerAspValTrp 1006

```

Db 620 TTTTGGTATGCTCCAGAAATGTTTAAATCAATCAAAATTTTATATGCTCCTGACGCTGG 679
QY 1007 SerPheGlyValValLeuTyrGluLeuPheThrTyrCysaspLysSerCysSerProSer 1026
Db 680 TCTTTTGGAGTCACCTCTGCATGAGCTGCTGACTTACTGTGATTCAGATTCTAGTCCCATG 739
QY 1027 AlaGluPheLeuArgMetMetGlyProGluArgGluGlyProProLeuCysArgLeuLeu 1046
Db 740 GCTTTGTTCTGAAATGATAGGCCCAACCCATGGCCAGATGACAGTCACAAGACTTGTG 799
QY 1047 GluLeuLeuAlaGluGlyArgLeuProProProProThrCysProThrGluValGln 1066
Db 800 AATACGTTAAAGAGGAAACGCTGCGTGCCACCTAACTGTCAGATGAGGTTTAT 859
QY 1067 GluLeuMetGlnLeuCysTrpAlaProGluProHisAspArgProAlaPheAlaThrLeu 1086
Db 860 CAACTTATGAGGAAATGCTGGGAATTCCAACCATCAATCGGACAGCTTTCAGAACCTT 919

```

Search completed: February 4, 2004, 11:29:00  
 Job time : 1207.04 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 4, 2004, 11:29:09 ; Search time 142 Seconds

(without alignments)  
3416.057 Million cell updates/sec

Title: US-09-397-967A-16

Perfect score: 1099

Sequence: 1 MAPSEETPLIPQSCSLSS.....RPAFATLSPQLDPLWRGPG 1099

# Scoring table:

OLIGO  
-Q=/cgn2\_6/ptodata/1/ina/5A.COMB.seq:  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delopt 6.0 , Delopt 7.0

Searched: 569978 seqs, 220691566 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1135299

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_6/ptodata/1/ina/5A.COMB.seq:  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=olip2n.rni -MINWATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09397967@cgn 1 71 @runat 03022004 175717 28949 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:  
5: /cgn2\_6/ptodata/1/ina/PTCUTS.COMB.seq:  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	180	16.4	4016	5	PCT-US95-08354A-1
2	62	5.6	3620	4	US-09-016-434-1055
3	62	5.6	3807	1	US-08-357-598-1
4	62	5.6	3807	2	US-09-003-289-1
5	62	5.6	3807	5	PCT-US95-16435-1
6	16	1.5	3435	3	US-09-046-158A-21
7	16	1.5	3495	1	US-08-446-038B-2
8	16	1.5	3495	1	US-08-446-0108-2
9	16	1.5	3495	1	US-08-805-445-2
10	16	1.5	3495	2	US-08-064-067D-2
11	16	1.5	3495	2	US-09-066-208-2
12	16	1.5	3495	4	US-08-980-080-3

13	1.5	3629	1	US-08-097-997A-8	Sequence 8, Appli
14	1.5	3629	3	US-08-665-574C-8	Sequence 8, Appli
15	1.5	3629	3	US-08-946-994-8	Sequence 8, Appli
16	1.5	4078	4	US-09-016-434-297	Sequence 297, App
17	1.5	4482	2	US-08-567-508C-1	Sequence 1, Appli
18	1.5	4482	3	US-09-196-480-1	Sequence 1, Appli
19	1.5	5117	4	US-09-972-800A-15	Sequence 15, Appli
20	1.2	41	1	US-08-481-003-9	Sequence 9, Appli
21	1.2	41	3	US-08-485-598-9	Sequence 9, Appli
22	1.2	45	1	US-08-481-003-8	Sequence 8, Appli
23	1.2	45	3	US-08-485-598-8	Sequence 8, Appli
24	1.1	151	1	US-08-222-616-11	Sequence 11, Appli
25	1.1	151	4	US-08-446-648-11	Sequence 11, Appli
26	1.1	151	5	PCT-US95-04228-11	Sequence 11, Appli
27	1.1	522	4	US-09-698-505A-29	Sequence 29, Appli
28	1.1	738	2	US-08-604-989A-8	Sequence 8, Appli
29	1.1	927	4	US-10-004-542-3	Sequence 3, Appli
30	1.1	1398	2	US-08-604-989A-9	Sequence 9, Appli
31	1.1	1521	2	US-08-604-989A-10	Sequence 10, Appli
32	1.1	1713	4	US-09-741-154-1	Sequence 1, Appli
33	1.1	1942	2	US-08-604-989A-11	Sequence 11, Appli
34	1.1	1987	2	US-08-876-882-1	Sequence 1, Appli
35	1.1	1987	4	US-09-315-928-1	Sequence 1, Appli
36	1.1	2000	4	US-08-426-509A-1	Sequence 1, Appli
37	1.1	2000	4	US-08-232-545-1	Sequence 1, Appli
38	1.1	2000	5	PCT-US95-05008-1	Sequence 1, Appli
39	1.1	2820	1	US-08-162-809-5	Sequence 5, Appli
40	1.1	2829	4	US-10-004-542-1	Sequence 1, Appli
41	1.1	2962	2	US-08-449-645A-10	Sequence 10, Appli
42	1.1	2962	2	US-08-702-367A-10	Sequence 10, Appli
43	1.1	2962	5	PCT-US95-04681-10	Sequence 10, Appli
44	1.1	2982	1	US-08-348-143-2	Sequence 2, Appli
45	1.1	2982	1	US-08-571-785-2	Sequence 2, Appli

## ALIGNMENTS

### RESULT 1

PCT-US95-08354A-1  
; Sequence 1, Application PC/TUS9508354A  
; GENERAL INFORMATION:  
; APPLICANT: Temple University - Of The  
; APPLICANT: Commonwealth System of Higher Education  
; TITLE OF INVENTION: JAK3 PROTEIN TYROSINE  
; TITLE OF INVENTION: KINASE AND DNA ENCODING THE SAME  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seidel, Gonda, Lavorgna  
; ADDRESSEE: & Monaco, P.C.  
; STREET: Suite 1800, Two Penn Center  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/08354A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/272,368  
; FILING DATE: 8 July 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 6056-203 PC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4016 nucleotides  
 TYPE: nucleic acid  
 STRANDEDNESS: single stranded  
 TOPOLOGY: linear  
 PCT-US95-08354A-1

Alignment Scores:  
 Pred. No.: 9, 6e-154 Length: 4016  
 Score: 180.00 Matches: 529  
 Percent Similarity: 97.24% Conservative: 0  
 Best Local Similarity: 97.24% Mismatches: 9  
 Query Match: 16.38% Indels: 15  
 Gaps: 0

US-09-397-967A-16 (1-1099) x PCT-US95-08354A-1 (1-4016)

QY	535	ArgArgGluValValAspGlyGluThrHisAspSerGluValLeuLeuLysValMetAsp	554
DB	2257	AGGCGGAGGTCTGGATGGTGAGACATGACTGGAAAGTCTCTGAGAGTCATGGAC	2316
QY	555	SerArgHisArgAsnCysMetGluSerPheLeuGluAlaAlaSerLeuMetSerGlnVal	574
DB	2317	TCAGACATCGGAAGTCTGAGTCTTTCTGGAAGCGCAAGCTTGATGAGCCAAAGTA	2376
QY	575	SerTyrProHisLeuValLeuLeuHisGlyValCysMetAlaGlyAspSerIleMetVal	594
DB	2377	TCTACCCGACCTGGTGTACTGACGCGCTGTCATGGTGAGACATCATGGTG	2436
QY	595	GlnGluPheValTyrLeuGlyAlaIleAspMetTyrLeuArgLysArgGlyHisLeuVal	614
DB	2437	CAGGAATTTGTGTATCTAGGAGCAATGTGCACTGCTGCGCAAGCGTGCCACCTGGTG	2496
QY	615	SerAlaSerTrpLysLeuGlnValThrLysGlnLeuAlaIleAlaLeuAsnTyrLeuGlu	634
DB	2497	TCAGCCAGCTGGAACCTGACAGGTGACCAAGCAGCTGGCATATGCCCTTAATCTGGAG	2556
QY	635	AspLysGlyLeuProHisGlyAsnValSerAlaArgLysValLeuLeuAlaArgGluGly	654
DB	2557	GACAAAGGCTCTCTCACGGCAAGCTCTACAGCAGGAAGGTGCTGCTGGCTGAGGGG	2616
QY	655	GlyAspGlyAsnProPheIleLysLeuSerAspProGlyValSerProThrValLeu	674
DB	2617	GGTGATGGGAATCCACCTTTCAATAGCTGAGTATCTGCTGCTGCTGCTGCTGCTGCTG	2676
QY	675	SerLeuGluMetLeuThrAspArgGlyProTrpValAlaProGluCysLeuGlnGluAla	694
DB	2677	AGCTGGAAATGCTCACCGACAGAAATACCTGGGTGGCCCCCGAAATGCTCCAGGAGGT	2736
QY	695	GlnThrLeuCysLeuGluAlaAspLysTrpGlyPheGlyAlaThrTrpGluValPhe	714
DB	2737	CAGACATCTGCTGGAGGCTGACAAGTGGGGCTTTGGAGCCACCGTGGAGGTGT-	2795
QY	715	GlnArg-GlyProAlaHisIleThrSerLeuGluProAlaLysLysLeuLysPheTyrGln	734
DB	2796	CAGCGGGGACCCCGCCACATCACTGCTGGAGCCCGCCCAAAAGCTGAAGTTCATGA	2855
QY	734	uAspGlnGlyGlnLeuProAlaLeuLysTrpThrGluLeuAlaGlyLeuIleThrGlnCys	754
DB	2856	GGACCGGACAGCTGCCGCTCTCAATGGACAGAACTGGCGGAGCTTATCACACAGTG	2915
QY	754	sMetAlaTyrAspProGlyArgArgProSerPheArgAlaIleLeuArgAspLeuAsnGln	774
DB	2916	CATGGCGTATGATCTGGCGCGCGCTCTCTCGAGCTATCTCCAGAGACCTCAACGG	2975
QY	774	YLeuIleThrSerAspTyrGluLeuLeuSerAspProThrProGlyIleProSerProArg	794
DB	2976	CCTCATTCATCAGATTACAGCTCTCTCA-GACCCACACCTGGCATCCCGAGTCTCTCG	3034
QY	794	gAspGluLeuCysValAlaGlyAlaGlnLeuTyrAlaCysGlnAspProAlaIlePheGln	814
DB	3035	AGATGAGCTGTGCG--T--GGCGCCAGCTCTATGCTGCCAGGACCCCGCCATATC-GA	3089

QY	814	uGluArgHisLeuLysTyrIleSerLeuLeuGlyLysGlyAsnPheGlySerValGluLeu	834
DB	3090	GGAGAGACACCTTAAGTACATCTCTTCTGGCAAGGCGCAACTTTGGCAGCGTGAGCT	3149
QY	834	uCysArgTyrAspProLeuGlyAspAsnThrGlyProLeuValAlaValLysGlnLeuGln	854
DB	3150	GTGCCGCTATGACCCCC--TG-GACAATACGGGACCCCTGGTGGCAGTGAAACACACTACA	3206
QY	854	nHisSerVal-ProAspGlnGlnArgAspPheGlnArgGluIleGlnLeuLysAlaL	874
DB	3207	GCACAGCGG-GCCAGACCCAGCAGAGGAGCTTCCACGGGAGATTCCAGATCCTTAAGGCTC	3265
QY	874	euHisSerAspPheIleValLysTyrArgGlyValSerTyrGlyProGlyArgGlnSerL	894
DB	3266	TGCACAGCGACTTCAAGTACCGGGGAGTCAAGCTATGGGCGAGCTGCCAGAGCC	3325
QY	894	euArgLeuValMetGluTyrIleLeuProSerGlyCysLeuArgAspLeu-LeuGlnArgHis	913
DB	3326	TGCGGTTGGTGTAGTACCTGCGGCGGCTGCTGCGAGA-CTTCTGCGAGCGCAT	3384
QY	914	--ArgGly-LeuHisThrAspArgLeuLeuPheAlaTrpGlnIleCysLysGlyMet	932
DB	3385	CGCGGGCCCTGCACACCGACCGCTACTGCTGCTTGGCTTGGCAGATCTGCAAGGCGATG	3444
QY	933	GluTyrLeuGlyAlaArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuVal	952
DB	3445	GAGTACCTGGTGGCGCGCTGCGGTACACCGTACCTGCTGCGCGCAACATCTTGGTG	3504
QY	953	GluSerGluAlaHisValLysIleAlaAspPheGlyLeuAlaLysLeuLeuProLeuGly	972
DB	3505	GAGAGCGAGGCTCATGTGAAGATCGCGGACTTCCGGCTCGCTAAGCTGCTGCCCTGGGA	3564
QY	973	LysAspTyrTyrValValArgGluProGlyGlnSerProIlePheTrpTyrAlaProGlu	992
DB	3565	AGGACTACTACGTGTCGCGAGCTCGCCAAAGCCCATCTTTTGGTATGATGCCGAG	3624
QY	993	SerLeuSerAspAsnIlePheSerArgGlnSerAspValTrpSerPheGlyValValLeu	1012
DB	3625	TCCCTATCTGACAAATCTTCTCCGCCAATCTGAGCTGTGGAGCTTGGAGTGGTGTG	3684
QY	1013	TyrGluLeuPheThrTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMet	1032
DB	3685	TACGAGCTCTTCACTACTGCGACAGAGCTGCGGCCCATCCCGTGAATCTTCTGGCATG	3744
QY	1033	MetGlyProGluArgGluGlyProProLeuCysArgLeuLeuGluLeuAlaGluGly	1052
DB	3745	ATGGGCGCTGAGCGTGNAGGACCCCGCTCTGCGGCTCTCTGGAGCTGCTGGCAGAGGC	3804
QY	1053	ArgArgLeuProProProProThrCysProThrGluValGlnGluLeuMetGlnLeuCys	1072
DB	3805	CGACGCTCCACACACCTCCGACCTGCCCGCCAGGTTCCAGGAGCTCATGCACTGTGC	3864

RESULT 2

US-09-016-434-1055  
 ; Sequence 1055, Application US/09016434  
 ; Patent No. 6500938

GENERAL INFORMATION:

APPLICANT: Janice Au-Young  
 APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 NUMBER OF SEQUENCES: 1490  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMEUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1055:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3620 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g1039418  
US-09-016-434-1055

Alignment Scores:  
Pred. No.: 1.47e-46 Length: 3620  
Score: 62.00 Matches: 96  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 5.64% Indels: 2  
DB: 4 Gaps: 0  
US-09-397-967A-16 (1-1099) x US-09-016-434-1055 (1-3620)

Qy 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnLeuValGluSerGluAlaHis 957  
Db 2922 CGCGCTGCGTGACCGCGACTGGCGCCGCGAAACATCTCTGTGGAGAGCGAGCAC 2981  
Qy 958 VallysileAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977  
Db 2982 GTCAAGATCGTGAATTCGGCTAGCTAAGCTGCTGGCGTTGAC-AAAGACTACTAGT 3040  
Qy 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997  
Db 3041 GGTCCGCGAGCCAGCCAGAGCCCAATTTCTGTATGCCCCGAATCCCTCTCGGCAA 3100  
Qy 997 nilePheSerArgGlnSerAspValTyrPsePheGlyValValLeuTyrGluLeuPheTh 1017  
Db 3101 CATCTTCTCTCGCAGTCAGACGCTCGAGCTTCGGGCTGCTCTGTACGAGCTCTTCAC 3160  
Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034  
Db 3161 CTACTCGCAAAAGCTGCAGCCCTCGCGCGAGTTCTCTCGGATGATGGA 3212

RESULT 3  
US-08-357-598-1  
Sequence 1, Application US/08357598  
Patent No. 5705625  
GENERAL INFORMATION:  
APPLICANT: Civin, Curt I.  
APPLICANT: Small, Donald  
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA

ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,598  
FILING DATE: 15-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-357-598-1

Alignment Scores:  
Pred. No.: 1.54e-46 Length: 3807  
Score: 62.00 Matches: 96  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 5.64% Indels: 2  
DB: 1 Gaps: 0  
US-09-397-967A-16 (1-1099) x US-08-357-598-1 (1-3807)

Qy 938 ArgArgCysValHisArgAspLeuAlaAlaArgAsnLeuValGluSerGluAlaHis 957  
Db 2985 CGCGCTGCGTGACCGCGACTGGCGCCGCGAAACATCTCTGTGGAGAGCGAGCAC 3044  
Qy 958 VallysileAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrVa 977  
Db 3045 GTCAAGATCGTGAATTCGGCTAGCTAAGCTGCTGGCGTTGAC-AAAGACTACTAGT 3103  
Qy 977 lValArgGluProGlyGlnSerProIlePheTyrAlaProGluSerLeuSerAspAs 997  
Db 3104 GGTCCGCGAGCCAGCCAGAGCCCAATTTCTGTATGCCCCGAATCCCTCTCGGCAA 3163  
Qy 997 nilePheSerArgGlnSerAspValTyrPsePheGlyValValLeuTyrGluLeuPheTh 1017  
Db 3164 CATCTTCTCTCGCAGTCAGACGCTTCGGGCTGCTCTGTACGAGCTCTTCAC 3223  
Qy 1017 rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly 1034  
Db 3224 CTACTCGCAAAAGCTGCAGCCCTCGCGCGAGTTCTCTCGGATGATGGA 3275

RESULT 4  
US-09-003-289-1  
Sequence 1, Application US/09003289  
Patent No. 5916792  
GENERAL INFORMATION:  
APPLICANT: Civin, Curt I.  
APPLICANT: Small, Donald  
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:

Thu Feb 5 11:06:33 2004

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/003,289  
FILING DATE: 15-DEC-1994  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/357,598  
FILING DATE: 15-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-09-003-289-1

Alignment Scores:  
Pred. No.: 1.54e-46 Length: 3807  
Score: 62.00 Matches: 96  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 5.64% Indels: 2  
DB: 2 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-003-289-1 (1-3807)

QY	938	ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis	957
DB	2985	CGCCGCTGCTGACCTTCGCGCTAGCTAAGCTGCTGCGCTTGCAC-AAAGACTACTACGT	3103
QY	958	ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa	977
DB	3045	GTCAGATCGTCTGACCTTCGCGCTAGCTAAGCTGCTGCGCTTGCAC-AAAGACTACTACGT	3103
QY	977	lValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs	997
DB	3104	GGTCCGCGAGCCAGGCGCAGAGCCCAATTTCTGGTATGCCCGCGAATCCCTCTCGGACAA	3163
QY	997	nilePheSerArgGlnSerAspValTyrPheTyrTyrAlaProGluSerLeuSerAspAs	997
DB	3164	CATCTTCTCTGCGCAGTCAGACGCTCGAGCTTCGGGGTCTGCTGACGAGCTCTTCAC	3223
QY	1017	rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly	1034
DB	3224	CTACTGCGACAAAGAGCTGCAGCCCTCGCGCGAGTTCTCGGGATGATGGGA	3275

RESULT 5  
PCT-US95-16435-1  
Sequence 1, Application PC/TUS9516435  
GENERAL INFORMATION:  
APPLICANT: The Johns Hopkins University School of Medicine  
TITLE OF INVENTION: NOVEL PROTEIN TYROSINE KINASE, JAK3  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16435  
FILING DATE: 15-DEC-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/033001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3807 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
PCT-US95-16435-1

Alignment Scores:  
Pred. No.: 1.54e-46 Length: 3807  
Score: 62.00 Matches: 96  
Percent Similarity: 97.96% Conservative: 0  
Best Local Similarity: 97.96% Mismatches: 1  
Query Match: 5.64% Indels: 2  
DB: 5 Gaps: 0

US-09-397-967A-16 (1-1099) x PCT-US95-16435-1 (1-3807)

QY	938	ArgArgCysValHisArgAspLeuAlaAlaArgAsnIleLeuValGluSerGluAlaHis	957
DB	2985	CGCCGCTGCTGACCTTCGCGCTAGCTAAGCTGCTGCGCTTGCAC-AAAGACTACTACGT	3103
QY	958	ValLysIleAlaAspPheGlyLeuAlaLysLeuLeuPro-LeuGlyLysAspTyrTyrVa	977
DB	3045	GTCAGATCGTCTGACCTTCGCGCTAGCTAAGCTGCTGCGCTTGCAC-AAAGACTACTACGT	3103
QY	977	lValArgGluProGlyGlnSerProIlePheTyrTyrAlaProGluSerLeuSerAspAs	997
DB	3104	GGTCCGCGAGCCAGGCGCAGAGCCCAATTTCTGGTATGCCCGCGAATCCCTCTCGGACAA	3163
QY	997	nilePheSerArgGlnSerAspValTyrPheTyrTyrAlaProGluSerLeuSerAspAs	997
DB	3164	CATCTTCTCTGCGCAGTCAGACGCTTCGGGGTCTGCTGACGAGCTCTTCAC	3223
QY	1017	rTyrCysAspLysSerCysSerProSerAlaGluPheLeuArgMetMetGly	1034
DB	3224	CTACTGCGACAAAGAGCTGCAGCCCTCGCGCGAGTTCTCGGGATGATGGGA	3275

RESULT 6  
US-09-046-158A-21  
Sequence 21, Application US/09046158A  
Patent No. 6187552  
GENERAL INFORMATION:  
APPLICANT: Roberts, Steven L.  
APPLICANT: Kaytes, Paul S.  
TITLE OF INVENTION: METHOD FOR IDENTIFYING INHIBITORS OF  
CYTOKINE RECEPTOR BINDING  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pharmacia & Upjohn Co., Intellectual Property  
ADDRESSEE: Legal Services  
STREET: 301 Henrietta Street  
CITY: Kalamazoo  
STATE: MI  
COUNTRY: USA  
ZIP: 49001  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk



COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,158A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Darinley Jr., James D.  
REGISTRATION NUMBER: 33,673  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 616/833-2210  
TELEFAX: 616/833-8897  
TELEX: 224401  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3435 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-046-158A-21

Alignment Scores:  
Pred. No.: 8.83e-05 Length: 3435  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 3 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-046-158A-21 (1-3435)  
QY 1003 SerAspValTrpSerPheGlyValValLeuTy-GluLeuPheThrTyr 1018  
Db 3106 TCAGATGTTGGAGCTTTGGAGTGGTCTGTATGAACTTTTCACATAC 3153

RESULT 7  
US-08-446-038B-2  
Sequence 2, Application US/08446038B  
Patent No. 5658791  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;  
APPLICANT: Harpur, Ailsa  
TITLE OF INVENTION: No. 5658791el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,038B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-Jun-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5658791-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5658791-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-No. 5658791-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5658791man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-08-446-038B-2

Alignment Scores:  
Pred. No.: 8.97e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 1 Gaps: 0

US-09-397-967A-16 (1-1099) x US-08-446-038B-2 (1-3495)  
QY 1003 SerAspValTrpSerPheGlyValValLeuTy-GluLeuPheThrTyr 1018  
Db 2695 TCAGATGTTGGAGCTTTGGAGTGGTCTGTATGAACTTTTCACATAC 2742

RESULT 8  
US-08-446-010B-2  
Sequence 2, Application US/08446010B  
Patent No. 5716818  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemiecki, Andrew;  
APPLICANT: Harpur, Ailsa  
TITLE OF INVENTION: No. 5716818el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,010B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 433  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,038  
FILING DATE: 19-MAY-1995  
APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-Jun-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-No. 5716818-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-No. 5716818-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian 88229/91  
FILING DATE: 27-No. 5716818-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Baer, Madeline F.  
REGISTRATION NUMBER: 36,437

REFERENCE/DOCKET NUMBER: LUD 5244.3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-08-446-010B-2

Alignment Scores:  
Pred. No.: 8.97e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
Gaps: 1

US-09-397-967A-16 (1-1099) x US-08-446-010B-2 (1-3495)

QY 1003 SerAspValTprSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
DB 2695 TCAGATGTGTGGAGCTTTGGAGTGGTCTATACGAACCTTTTCACATAC 2742

RESULT 9  
US-08-805-445-2  
Sequence 2, Application US/08805445  
Patent No. 5821069  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;  
APPLICANT: Harpur, Ailsa  
TITLE OF INVENTION: No. 5821069el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/805,445  
FILING DATE: 25-FEB-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/446,038  
FILING DATE: 19-MAY-1995  
APPLICATION NUMBER: 08/064,067  
FILING DATE: 30-JUN-1993  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-NO. 5821069-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-NO. 5821069-1990  
APPLICATION DATA: Australian 88229/91  
FILING DATE: 27-NO. 5821069-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5821069man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-08-805-445-2

Alignment Scores:  
Pred. No.: 8.97e-05 Length: 3495  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
Gaps: 1

US-09-397-967A-16 (1-1099) x US-08-805-445-2 (1-3495)

QY 1003 SerAspValTprSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018  
DB 2695 TCAGATGTGTGGAGCTTTGGAGTGGTCTATACGAACCTTTTCACATAC 2742

RESULT 10  
US-08-064-067D-2  
Sequence 2, Application US/08064067D  
Patent No. 5852184  
GENERAL INFORMATION:  
APPLICANT: Wilks, Andrew F.; Ziemiacki, Andrew;  
APPLICANT: Harpur, Ailsa  
TITLE OF INVENTION: No. 5852184el Protein Tyrosine Kinase  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/064,067D  
FILING DATE: 30-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08889  
FILING DATE: 26-NO. 5852184-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: Australian PK3594/90  
FILING DATE: 28-NO. 5852184-1990  
APPLICATION DATA: Australian 88229/91  
FILING DATE: 27-NO. 5852184-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 5852184man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5244  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-688-9200  
TELEFAX: 212-838-3884  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3495 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: nucleic acid  
US-08-064-067D-2

Alignment Scores:  
Pred. No.: 8.97e-05 Length: 3495

```

Query Match: 1.46% Indels: 0
DB: 2 Gaps: 0

US-09-397-967A-16 (1-1099) x US-09-066-208-2 (1-3495)

Qy 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 2695 TCAGATGTGGAGCTTTGGAGTGGTTCTATACGAACCTTTTCACATAC 2742

RESULT 12
US-08-980-080-3
; Sequence 3, Application US/08980080
; Patent No. 6312941
; GENERAL INFORMATION:
; APPLICANT: CARTER-SU, CHRISTIN
; APPLICANT: RUI, LIANG-YOU
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: SIGNALING PATHWAY AGONISTS AND ANTAGONISTS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 MONTGOMERY STREET, SUITE 2200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,080
; FILING DATE: 26-NOV-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CARROLL, PETER G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UM-03036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3495 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2979
US-08-980-080-3

Alignment Scores:
Pred. No.: 8.97e-05 Length: 3495
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.46% Indels: 0
DB: 4 Gaps: 0

US-09-397-967A-16 (1-1099) x US-08-980-080-3 (1-3495)

Qy 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
|||||
Db 2695 TCAGATGTGGAGCTTTGGAGTGGTTCTATACGAACCTTTTCACATAC 2742

RESULT 13
US-08-097-997A-8
; Sequence 8, Application US/08097997A
; Patent No. 5728536
; GENERAL INFORMATION:

```

```

; APPLICANT: Ihle, James N.
; APPLICANT: Silvennoinen, Ollie
; APPLICANT: Witthuhn, Bruce A.
; APPLICANT: Quelle, Frederick W.
; TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
; TITLE OF INVENTION: Transduction
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,997A
; FILING DATE: 29-JULY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fox, Samuel L.
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0656.0370000/SLF/GKT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3629 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 94...3480
;
; US-08-097-997A-8
;
; Alignment Scores:
; Pred. No.: 9,29e-05 Length: 3629
; Score: 16.00 Matches: 16
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 1.46% Indels: 0
; DB: 1 Gaps: 0
;
; US-09-397-967A-16 (1-1099) x US-08-097-997A-8 (1-3629)
;
; QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
; DB 3196 TCAGATGTGGAGCTTTGGAGTGTCTATACGAACCTTTTCACATAC 3243
;
; RESULT 14
; US-08-665-574C-8
; Sequence 8, Application US/08665574C
; Patent No. 6136595
; GENERAL INFORMATION:
; APPLICANT: Ihle, James N.
; APPLICANT: Silvennoinen, Ollie
; APPLICANT: Witthuhn, Bruce A.
; TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine
; TITLE OF INVENTION: Signal Transduction
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/097,997A
; FILING DATE: 29-JULY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fox, Samuel L.
; REGISTRATION NUMBER: 30,353
; REFERENCE/DOCKET NUMBER: 0656.0370000/SLF/GKT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3629 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 94...3480
;
; US-08-097-997A-8
;
; Alignment Scores:
; Pred. No.: 9,29e-05 Length: 3629
; Score: 16.00 Matches: 16
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 1.46% Indels: 0
; DB: 1 Gaps: 0
;
; US-09-397-967A-16 (1-1099) x US-08-097-997A-8 (1-3629)
;
; QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
; DB 3196 TCAGATGTGGAGCTTTGGAGTGTCTATACGAACCTTTTCACATAC 3243
;
; RESULT 15
; US-08-946-994-8
; Sequence 8, Application US/08946994
; Patent No. 6210654
; GENERAL INFORMATION:
; APPLICANT: Ihle, James N.
; APPLICANT: Silvennoinen, Ollie
; APPLICANT: Witthuhn, Bruce A.
; APPLICANT: Quelle, Frederick W.
; TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
; TITLE OF INVENTION: Transduction
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,574C
; FILING DATE: 18-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/282,012
; FILING DATE: 29-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/097,997
; FILING DATE: 29-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/118,968
; FILING DATE: 09-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Bugalsky, Lawrence B.
; REGISTRATION NUMBER: 35,086
; REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/LBB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3629 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 94...3480
;
; US-08-665-574C-8
;
; Alignment Scores:
; Pred. No.: 9,29e-05 Length: 3629
; Score: 16.00 Matches: 16
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 1.46% Indels: 0
; DB: 3 Gaps: 0
;
; US-09-397-967A-16 (1-1099) x US-08-665-574C-8 (1-3629)
;
; QY 1003 SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr 1018
; DB 3196 TCAGATGTGGAGCTTTGGAGTGTCTATACGAACCTTTTCACATAC 3243
;
; RESULT 15
; US-08-946-994-8
; Sequence 8, Application US/08946994
; Patent No. 6210654
; GENERAL INFORMATION:
; APPLICANT: Ihle, James N.
; APPLICANT: Silvennoinen, Ollie
; APPLICANT: Witthuhn, Bruce A.
; APPLICANT: Quelle, Frederick W.
; TITLE OF INVENTION: Jak Kinases and Regulation of Cytokine Signal
; TITLE OF INVENTION: Transduction
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/946,994  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/665,574  
FILING DATE: 18-JUN-1996  
APPLICATION NUMBER: 08/282,012  
FILING DATE: 29-JUL-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/097,997  
FILING DATE: 29-JUL-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/118,968  
FILING DATE: 09-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Fox, Samuel L.  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0656.0370002/SLF/GKT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3629 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 94..3480  
US-08-946-994-8

Alignment Scores:  
Pred. No.: 9.29e-05 Length: 3629  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.46% Indels: 0  
DB: 3 Gaps: 0

US-09-397-967A-16 (1-1099) x US-08-946-994-8 (1-3629)

Cy	1003	SerAspValTrpSerPheGlyValValLeuTyrGluLeuPheThrTyr	1018
Db	3196	TCAGATGTTGGACCTTTGGAGTGGTCTATACGAACCTTTTCACATAC	3243

Search completed: February 4, 2004, 17:52:55  
Job time : 187 secs

